Iwaniec's Conjecture on The Beurling-Ahlfors Transform

Zoe Nieraeth
3756351

Universiteit Utrecht

Master's thesis Mathematical Sciences

Supervisor:
Dr.ir. M.C. Veraar
Examiners:
Prof. Dr. E.P. van den Ban
Prof. Dr. S.M. Verduyn Lunel

July 30, 2016

Abstract

Inspired by Astala, Iwaniec, Prause and Saksman's partial result of Morrey's problem regarding rank-one convex and quasiconvex functions on the functionals from Burkholder's martingale theory, we discuss and relate several open problems in different fields of mathematics. In particular, we discuss the theory of Calderón and Zygmund regarding the L^p-boundedness of the Beurling-Ahlfors transform $\mathscr{B}$ for $1 < p < \infty$ to formulate Iwaniec's conjecture regarding the precise operator norms of $\mathscr{B}$. Moreover, we discuss its consequences in the theory of quasiconformal mappings. Finally, we discuss the notions of rank-one convexity and quasiconvexity, motivated by their role in the theory of calculus of variations, and show how a positive answer to Morrey's conjecture implies quasiconvexity of the Burkholder functional, which, in turn, is shown to imply Iwaniec's conjecture.

Preface

This thesis was written under supervision of Mark Veraar of the Analysis Group at the Delft University of Technology as part of the master Mathematical Sciences at Utrecht University.

Acknowledgements

First and foremost I would like to thank my supervisor Mark Veraar, not only for introducing me to this subject, but also for providing me with a working environment at the Delft University of Technology that inspires academic excellence. Not only was I able to obtain direct feedback to my ideas from Mark and the other people in the analysis department, but I was also able to learn about current research being done there through various seminars and lectures organized by the staff to provide me with a broader mathematical perspective.

I wish to express my gratitude to my examiners Erik van den Ban and Sjoerd Verduyn Lunel for their enthusiasm about the thesis and for taking their time to listen to and read about my work.

I would like to thank Nick Lindemulder for providing commentary on the earlier drafts of this thesis, for helping me work out some of my ideas, and for providing helpful suggestions to enhance the quality of the text.

I am indebted to Adriána Szilágiová, who provided numerous comments leading to improvement in the formulation and readability of the thesis, all while she was dealing with her own master's thesis and the pressure that comes with it.

Contents

Introduction

It was in a seminar on complex analysis in 1949 at the University of Uppsala where Arne Beurling introduced a two-dimensional analogue of the Hilbert transform, which we now call the Beurling-Ahlfors transform, and proved that it extends to an isometry of L^2 as a generalization of Hilbert's result for the Hilbert transform, see [Be, p. 460]. In 1955, this operator found its way into the theory of quasiconformal mappings when it was used by Lars Ahlfors to establish the existence of solutions to certain partial differential equations known as Beltrami equations, see [Al]. Notably, the mathematician Ilia Vekua had done work in this area earlier in the same year in [Ve]. Ahlfors was unaware of this fact and the results he found were independent of Vekua's results. It was Vekua's student Bogdan Bojarski who combined their efforts in 1957 in the seminal paper *Generalized Solutions of a System of First Order Differential Equations of Elliptic Type with Discontinuous Coefficients*, see [Bo].

Bojarski used the newly developed theory of Calderón and Zygmund on singular integral operators to the Beurling-Ahlfors transform to establish L^p estimates of solutions to Beltrami equations. Having these integrability results in mind, in 1982 the Polish mathematician Tadeusz Iwaniec published the article *Extremal Inequalities In Sobolev Spaces and Quasiconformal Mappings*, see [Iw], in which he conjectures precise values of the operator norm in L^p of the the Beurling-Ahlfors transform. In the same year, Donald Burkholder was independently working on his martingale theory which happened to feature the same values from Iwaniec's conjecture, see [Bu]. While Iwaniec's conjecture has yet to be settled, it has been through Burkholder's estimates that the most progress has been made. The study of the functionals obtained through Burkholder's theory relates back to notions of convexity introduced in the setting of calculus of variations in 1952 by Charles Morrey, see [Mo], in which there is an outstanding open problem, known as Morrey's conjecture, on relating the notions of quasiconvexity and rank-one convexity.

As a culmination of these ideas, Astala, Iwaniec, Prause and Saksman obtain a partial result with respect to Morrey's conjecture in 2010 in the article *Burkholder Integrals, Morrey's Problem and Quasiconformal Mappings*, see [AIPS]. We let this result inspire us to delve into the theory and to explore its history.

Goal and outline

The goal of this thesis is not to prove any new results, but to give an overview of the theory and ideas necessary to understand Iwaniec's conjecture and several related conjectures due to Burkholder's estimates.

The main text of the thesis is split into four sections. In Section 1 we provide some preliminary notions and results regarding L^p-spaces.

In Section 2 we first establish L^p-boundedness of the Hilbert transform and we establish its precise L^p-norms. Then we use Calderón and Zygmund's Method of Rotations to establish L^p-boundedness of the Riesz transforms and the Beurling-Ahlfors transform. We also establish a lower bound of the L^p-norms of the Beurling-Ahlfors transform.

In Section 3 we give an introduction into the theory of quasiconformal mappings. Moreover, we explain Iwaniec's reasoning on how he came to his conjecture.

In the last section, Section 4, we first describe how one of Burkholder's estimates can be used to deduce results regarding the operator norm of the Beurling-Ahlfors transform and we explain how this is related to the study of the Burkholder functional. We then give an introduction into the theory of calculus of variations and the related notions of quasiconvexity and rank-one convexity. This leads us to Morrey's conjecture on the equivalence of these convexity notions in two dimensions and to conjectures related to the quasiconvexity of the Burkholder functional. We conclude the section by giving an overview of the conjectures.

1 Preliminaries: L^p-Spaces

This section deals with some preliminary facts we will be using on L^p-spaces and some conventions we will be working with. Whenever we speak of a *function* we mean a map whose codomain is the field $\mathbf{C}$ of complex numbers. Naturally, all our function spaces will be vector spaces over $\mathbf{C}$. When we are working with functions defined on $\mathbf{R}^2$, we will use the standard identification $\mathbf{C} \cong \mathbf{R}^2$. We will usually denote the coordinates on $\mathbf{C}$ by $z = x + iy$. In an attempt to make our notation less cumbersome we will sometimes consider z, x, and y to be functions, where one might interpret z as the identity function on $\mathbf{C}$ and x and y as taking the respective real and imaginary parts of a complex number. It should be implied by the context when these letters refer to functions rather than values and vice-versa. When working in the Fourier domain we will usually denote the coordinates by $\zeta = \xi + i\eta$, working under similar conventions. Sometimes we will step away from these conventions when we wish to generalize to a setting on $\mathbf{R}^n$ for $n \in \mathbf{N}$. In this case we will denote the coordinates by $x = (x_1, \dots, x_n)$ or sometimes $y = (y_1, \dots, y_n)$. In the 1-dimensional case we will also sometimes use t.

When $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are normed (both real or both complex) vector spaces, then we define the (extended) operator norm by

$$\|\cdot\|_{\mathcal{L}(X,Y)} : \{\mathscr{L} : X \to Y \mid \mathscr{L} \text{ linear}\} \to [0, \infty], \quad \|\mathscr{L}\|_{\mathcal{L}(X,Y)} = \sup_{\substack{x \in X \\ \|x\|_X = 1}} \|\mathscr{L}x\|_Y,$$

and set

$$\mathcal{L}(X, Y) = \{\mathscr{L} : X \to Y \mid \mathscr{L} \text{ linear}, \|\mathscr{L}\|_{\mathcal{L}(X,Y)} < \infty\}.$$

We will also write $\mathcal{L}(X) := \mathcal{L}(X, X)$.

Throughout this section we let U be a non-empty open subset of $\mathbf{R}^n$. We denote by $L^0(U)$ the space of equivalence classes of Lebesgue measurable functions on U, where two functions are deemed equivalent if they are equal almost everywhere. We will commit the usual abuse of notation where we identify functions with their equivalence class, e.g., we will write $f \in L^0(U)$ for a function f rather than its corresponding equivalence class. For $p \in [1, \infty]$ we define the (extended) norms

$$\|\cdot\|_p : L^0(U) \to [0, \infty], \quad \|f\|_p := \begin{cases} \left(\int_U |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}} & \text{if } p \in [1, \infty[\\ \operatorname*{ess\,sup}_{x \in U} |f(x)| & \text{if } p = \infty, \end{cases}$$

where $\mathrm{d}x$ means integration with respect to the Lebesgue measure (where x represents the coordinates we are using) and where

$$\operatorname*{ess\,sup}_{x \in U} f(x) = \inf\{c \in \mathbf{R} \mid f(x) \leq c \text{ for a.e. } x \in U\}.$$

for real-valued $f \in L^0(U)$. We then set

$$L^p(U) := \{f \in L^0(U) \mid \|f\|_p < \infty\},$$

which are Banach spaces when equipped with their respective norms.

Let $p \in]1,\infty[$. We then call $p' := p/(p-1)$ the (Hölder) conjugate exponent of p. The map $q \mapsto q/(q-1)$ gives a bijection from $]1,2]$ to $[2,\infty[$ and vice-versa. If $f, g : U \to \mathbf{C}$ are measurable functions so that fg is integrable, then we write

$$\langle f, g \rangle := \int_U f(x)g(x)\,\mathrm{d}x.$$

The pairing $\langle \cdot, \cdot \rangle$ restricts to the dual pairing $L^p(U) \times L^{p'}(U) \to \mathbf{C}$, which is well-defined by Hölder's inequality. Then the maps

$$\begin{aligned} L^p(U) &\to (L^{p'}(U))^*, \quad f \mapsto \langle f, \cdot \rangle \\ L^{p'}(U) &\to (L^p(U))^*, \quad g \mapsto \langle \cdot, g \rangle \end{aligned}$$

are isometric linear isomorphisms.

Since the inclusion $\iota : C_c^\infty(U) \hookrightarrow L^{p'}(U)$ is continuous with dense range, see Appendix A, we find that the restriction map

$$(L^{p'}(U))^* \to \mathcal{D}'(U), \quad u \mapsto u|_{C_c^\infty(U)} = u \circ \iota$$

is a continuous injection. This allows us to give an alternative description of $L^p(U)$ by defining

$$\|\cdot\|_p : \mathcal{D}'(U) \to [0,\infty], \quad \|u\|_p = \sup_{\substack{\phi \in C_c^\infty(U) \\ \|\phi\|_{p'}=1}} |u(\phi)|$$

and setting

$$L^p(U) = \left\{ u \in \mathcal{D}'(U) \middle| \|u\|_p < \infty \right\}.$$

The first definition of $L^p(U)$ yields a space that is isometrically isomorphic to this new space through the map

$$f \mapsto \langle f, \cdot \rangle|_{C_c^\infty(U)}.$$

To see why this map is an isometry we will state a general lemma which will be used several times later on.

1.1 Lemma. *Let E be a normed vector space and F a Banach space. Suppose $V \subseteq E$ is a dense subspace, equipped with the restricted norm of E. Then the restriction map $\rho : \mathcal{L}(E,F) \to \mathcal{L}(V,F)$, $\rho(\mathscr{L}) := \mathscr{L}|_V$ is an isometric linear isomorphism.*

Proof. Note that ρ is linear. First we will show that ρ is isometric. Thus, we need to show that for any $\mathscr{L} \in \mathcal{L}(E,F)$ we have

$$\|\mathscr{L}\|_{\mathcal{L}(E,F)} = \|\mathscr{L}|_V\|_{\mathcal{L}(V,F)}. \tag{1.1}$$

The inequality

$$\|\mathscr{L}|_V\|_{\mathcal{L}(V,F)} = \sup_{\substack{x \in V \\ \|x\|_E = 1}} \|\mathscr{L}x\|_F \leq \|\mathscr{L}\|_{\mathcal{L}(E,F)}$$

is clear. For the converse inequality, note that any $x \in E$ with $\|x\|_E = 1$ can be approximated by a sequence $(x_j)_{j \in \mathbf{N}}$ in V such that $\|x_j\|_E = 1$ for all $j \in \mathbf{N}$. Then

$$\|\mathscr{L}x\|_F = \lim_{j \to \infty} \|\mathscr{L}x_j\|_F \leq \|\mathscr{L}|_V\|_{\mathcal{L}(V,F)}.$$

Hence, $\|\mathscr{L}\|_{\mathcal{L}(E,F)} \leq \|\mathscr{L}|_V\|_{\mathcal{L}(V,F)}$. This proves (1.1).

It remains to show that ρ is surjective. Let $\mathscr{M} \in \mathcal{L}(V, F)$. Then we have

$$\|\mathscr{M}x'\|_F \leq \|\mathscr{M}\|_{\mathcal{L}(V,F)}\|x'\|_E \quad \text{for all } x' \in V. \tag{1.2}$$

Let $x \in E$. Then there is a sequence $(x_j)_{j\in\mathbf{N}}$ in V that converges to x in E. Then this sequence is a Cauchy sequence in E. Hence, by taking $x' = x_j - x_k$ in (1.2) for $j, k \in \mathbf{N}$, we see that the sequence $(\mathscr{M}x_j)_{j\in\mathbf{N}}$ is a Cauchy sequence in F. Since F is complete, this means that there is some $y \in F$ so that $(\mathscr{M}x_j)_{j\in\mathbf{N}}$ converges to y. Note that if $(x'_j)_{j\in\mathbf{N}}$ is any other sequence in V that converges to x in E, then $(\mathscr{M}x'_j)_{j\in\mathbf{N}}$ is again convergent. By taking $x' = x'_j - x_j$ in (1.2) it follows that $(\mathscr{M}x'_j)_{j\in\mathbf{N}}$ must also converge to y.

Now we can define a map $\mathscr{L} : E \to F$ by setting $\mathscr{L}x := y$. This map is linear, and coincides with $\mathscr{M}$ on V. By another approximation argument using (1.2), it follows that $\mathscr{L}$ is bounded. We conclude that $\mathscr{L} \in \mathcal{L}(E, F)$ and $\rho(\mathscr{L}) = \mathscr{M}$. The assertion follows. □

To emphasize, a particular consequence of Lemma 1.1 is that

$$\|g\|_{p'} = \|\langle\cdot, g\rangle\|_{\mathcal{L}(L^p(\mathbf{R}^n),\mathbf{C})} = \sup_{\substack{\phi\in C_c^\infty(\mathbf{R}) \\ \|\phi\|_p=1}} |\langle\phi, g\rangle| \quad \text{for all } g \in L^{p'}(\mathbf{R}^n). \tag{1.3}$$

Sometimes we will tacitly use the second definition of $L^p(\mathbf{R}^n)$, which should be clear from the context.

We will be using the following basic result:

1.2 Lemma. *Let $\mathscr{L} \in \mathcal{L}(L^p(U))$. Then there is a unique dual operator $\mathscr{L}^* \in \mathcal{L}(L^{p'}(U))$ satisfying*

$$\langle\mathscr{L}f, g\rangle = \langle f, \mathscr{L}^*g\rangle$$

for all $f \in L^p(U)$ and $g \in L^{p'}(U)$. Moreover, this operator satisfies

$$\|\mathscr{L}^*\|_{\mathcal{L}(L^{p'}(U))} = \|\mathscr{L}\|_{\mathcal{L}(L^p(U))}.$$

Proof. For any $g \in L^{p'}(U)$ the map $L^p(U) \to \mathbf{C}$, $f \mapsto \langle\mathscr{L}f, g\rangle$ lies in $(L^p(U))^*$. Hence, there is a unique element $h \in L^{p'}(U)$ so that $\langle\mathscr{L}f, g\rangle = \langle f, h\rangle$ for all $f \in L^p(U)$. Setting $\mathscr{L}^*g := h$ yields a linear operator $\mathscr{L}^* : L^{p'}(U) \to L^{p'}(U)$ satisfying

$$\|\mathscr{L}^*\|_{\mathcal{L}(L^{p'}(U))} = \sup_{\substack{f\in L^p(U),\, g\in L^{p'}(U) \\ \|f\|_p=\|g\|_{p'}=1}} |\langle\mathscr{L}f, g\rangle| = \|\mathscr{L}\|_{\mathcal{L}(L^p(U))}.$$

The assertion follows. □

While we won't make much use of the additional Hilbert space structure of $L^2(U)$, we do want to remark the following:

1.3 Remark. The inner product on $L^2(U)$ is defined by $(f,g)_2 := \langle f, \overline{g}\rangle$. More generally, for $p \in]1,\infty[$ we can define $(f,g) := \langle f, \overline{g}\rangle$ for $f \in L^p(U)$, $g \in L^{p'}(U)$. One then finds that for every $\mathscr{L} \in \mathcal{L}(L^p(U))$ there is a unique conjugate transpose operator $\mathscr{L}^\dagger \in \mathcal{L}(L^{p'}(U))$ so that

$$(\mathscr{L}f, g) = (f, \mathscr{L}^\dagger g)$$

for all $f \in L^p(U)$, $g \in L^{p'}(U)$. This relates to the dual operator of $\mathscr{L}$ through the formula $\mathscr{L}^* f = \overline{\mathscr{L}^\dagger \overline{f}}$ and therefore also satisfies

$$\|\mathscr{L}^\dagger\|_{\mathcal{L}(L^{p'}(U))} = \|\mathscr{L}\|_{\mathcal{L}(L^p(U))}.$$

To illustrate the differences between the notion of the dual operator and the conjugate transpose operator we note that the Fourier transform $\mathscr{F}$ viewed as an operator in $\mathcal{L}(L^2(\mathbf{R}^n))$ now satisfies $\mathscr{F}^* = \mathscr{F}$, while $\mathscr{F}^\dagger = \mathscr{F}^{-1}$. $\diamondsuit$

We say that an open set $V \subseteq U$ is relatively compact in U, if $\overline{V} \subseteq U$ and $\overline{V}$ is compact. For $p \in [1,\infty]$ we define the local L^p-spaces

$$L^p_{loc}(U) = \{f \in L^0(U) \mid f|_V \in L^p(V) \text{ for all relatively compact } V \subseteq U\}.$$

We note that Hölder's inequality implies that $L^p_{loc}(U) \subseteq L^1_{loc}(U)$ for all $p \in [1,\infty]$.

1.4 Lemma. *Let $p \in [1,\infty]$ and $f \in L^0(U)$. Then $f \in L^p_{loc}(U)$ if and only if $\phi f \in L^p(U)$ for all $\phi \in C_c^\infty(U)$.*

Proof. We consider the cases where $p \in [1,\infty[$. The case $p = \infty$ is similar.

If $f \in L^p_{loc}(U)$ and $\phi \in C_c^\infty(U)$, then we can pick a relatively compact set $V \subseteq U$ so that $\operatorname{supp}\phi \subseteq V$. Then

$$\int_U |\phi(x)f(x)|^p \,\mathrm{d}x = \int_V |\phi(x)f(x)|^p \,\mathrm{d}x \le \|\phi\|_\infty^p \int_V |f(x)|^p \,\mathrm{d}x < \infty,$$

as desired.

For the converse, suppose $f \in L^0(U)$ satisfies $\phi f \in L^p_{loc}(U)$ for all $\phi \in C_c^\infty(U)$. Let $V \subseteq U$ be relatively compact and pick a cutoff function $\chi \in C_c^\infty(U)$ so that $\chi|_V = 1$. Then

$$\int_V |f(x)|^p \,\mathrm{d}x = \int_V |\chi(x)f(x)|^p \,\mathrm{d}x \le \int_U |\chi(x)f(x)|^p \,\mathrm{d}x < \infty.$$

The assertion follows. $\square$

In Section 3 we will be working in the Sobolev spaces

$$W^{1,p}(U) := \{f \in \mathcal{D}'(U) \mid f, \partial_j f \in L^p(U) \text{ for } j \in \{1,\dots,n\}\}$$
$$W^{1,p}_{loc}(U) := \{f \in \mathcal{D}'(U) \mid f, \partial_j f \in L^p_{loc}(U) \text{ for } j \in \{1,\dots,n\}\}$$

for $p \in [1,\infty]$. The norm $\|f\|_{W^{1,p}(U)} := \|f\|_p + \sum_{j=1}^n \|\partial_j f\|_p$ turns $W^{1,p}(U)$ into a Banach space.

We denote by $\mathbf{R}_+$ the strictly positive real numbers.

1.5 Theorem. *Let $p \in [1\,\infty[$. Then the space $C_c^\infty(\mathbf{R}^n)$ is dense in $W^{1,p}(\mathbf{R}^n)$.*

Proof. We denote by $W_c^{1,p}(\mathbf{R}^n)$ the space of those elements of $W^{1,p}(\mathbf{R}^n)$ that have compact support. The proof will be in two steps. First we will show that $W_c^{1,p}(\mathbf{R}^n)$ lies in $\overline{C_c^\infty(\mathbf{R}^n)}$, where the bar denotes taking the closure in $W^{1,p}(\mathbf{R}^n)$. Then we will show that $\overline{W_c^{1,p}(\mathbf{R}^n)} = W^{1,p}(\mathbf{R}^n)$.

For the first step, let $(\phi_\varepsilon)_{\varepsilon\in\mathbf{R}_+}$ in $C_c^\infty(\mathbf{R}^n)$ denote the standard mollifier, see Definition A.10. Let $\varepsilon \in \mathbf{R}_+$ and pick any $f \in W_c^{1,p}(\mathbf{R}^n)$. Then it follows from Lemma A.13 that $f * \phi_\varepsilon, \partial_j f * \phi_\varepsilon \in C_c^\infty(\mathbf{R}^n)$ for all $j \in \{1, \ldots, n\}$. Moreover, since $L^p(\mathbf{R}^n) \subseteq \mathcal{S}'(\mathbf{R}^n)$ and $C_c^\infty(\mathbf{R}^n) \subseteq \mathcal{S}(\mathbf{R}^n)$ it follows from Proposition B.41 that $\partial_j(f * \phi_\varepsilon) = \partial_j f * \phi$ for all $j \in \{1, \ldots, n\}$. It follows from Theorem A.7 that, taking limits in $L^p(\mathbf{R}^n)$, we have

$$\lim_{\varepsilon\downarrow 0} f * \phi_\varepsilon = f, \quad \lim_{\varepsilon\downarrow 0} \partial_j f * \phi_\varepsilon = \partial_j f \quad \text{for all } j \in \{1, \ldots, n\}.$$

This implies that $f * \phi_\varepsilon \to f$ in $W^{1,p}(\mathbf{R}^n)$ as $\varepsilon \downarrow 0$, proving that $W_c^{1,p}(\mathbf{R}^n) \subseteq \overline{C_c^\infty(\mathbf{R}^n)}$, as desired.

For the second step, we let $f \in W^{1,p}(\mathbf{R}^n)$ be arbitrary. Pick a cutoff function $\chi \in C_c^\infty(\mathbf{R}^n)$ satisfying $\chi(\mathbf{R}^n) \subseteq [0, 1]$, $\chi(x) = 1$ when $|x| < 1$. Then, for each $k \in \mathbf{N}$, we define $\chi_k \in C_c^\infty(\mathbf{R}^n)$ by $\chi_k(x) := \chi(x/k)$. Then $\chi_k f \in L^p(\mathbf{R}^n)$ and, by the Leibniz rule for differentiation of the product of a smooth function and a distribution, see Lemma B.14, we have

$$\partial_j(\chi_k f) = (\partial_j \chi_k)f + \chi_k \partial_j f \in L^p(\mathbf{R}^n) \tag{1.4}$$

for all $j \in \{1, \ldots, n\}$ so that $\chi_k f \in W_c^{1,p}(\mathbf{R}^n)$ for all $k \in \mathbf{N}$.

Since $\chi_k(x) = \chi(x/k) \to \chi(0) = 1$ as $k \to \infty$ for all $x \in \mathbf{R}^n$, we conclude from Lebesgue's Dominated Convergence Theorem that $\chi_k f \to f$ in $L^p(\mathbf{R}^n)$ as $k \to \infty$. Moreover, since $\partial_j \chi_k(x) = 0$ for $|x| < k$, a similar argument shows, using (1.4), that also $\partial_j(\chi_k f) \to \partial_j f$ in $L^p(\mathbf{R}^n)$ as $k \to \infty$ for all $j \in \{1, \ldots, n\}$. We conclude that $\chi_k f \to f$ in $W^{1,p}(\mathbf{R}^n)$ as $k \to \infty$. This proves that $\overline{W_c^{1,p}(\mathbf{R}^n)} = W^{1,p}(\mathbf{R}^n)$, as desired.

Finally, we observe that we have shown that

$$W^{1,p}(\mathbf{R}^n) = \overline{W_c^{1,p}(\mathbf{R}^n)} \subseteq \overline{C_c^\infty(\mathbf{R}^n)} \subseteq W^{1,p}(\mathbf{R}^n).$$

This proves the result. □

The following lemma characterizes local Sobolev spaces.

1.6 Lemma. *Let $p \in [1, \infty]$ and let $f \in \mathcal{D}'(U)$. Then $f \in W_{loc}^{1,p}(U)$ if and only if $\phi f \in W^{1,p}(U)$ for all $\phi \in C_c^\infty(U)$.*

Proof. Suppose $f \in W_{loc}^{1,p}(U)$. Let $\phi \in C_c^\infty(U)$ and $j \in \{1, \ldots, n\}$. Since $\partial_j f \in L_{loc}^p(U)$ and $\partial_j \phi \in C_c^\infty(U)$, Lemma 1.4 and the Leibniz rule for differentiation of the product of a smooth function and a distribution, see Lemma B.14, tell us that

$$\partial_j(\phi f) = (\partial_j \phi)f + \phi(\partial_j f) \in L^p(U), \quad \phi f \in L^p(U).$$

We conclude that $\phi f \in W^{1,p}(U)$.

For the converse, suppose $f \in \mathcal{D}'(U)$ satisfies $\phi f \in W^{1,p}(U)$ for all $\phi \in C_c^\infty(U)$. Then, in particular, $\phi f \in L^p(U)$ for all $\phi \in C_c^\infty(U)$ so that $f \in L^p_{loc}(U)$ by Lemma 1.4. Moreover, for $j \in \{1, \ldots, n\}$ we have

$$\phi \partial_j f = \partial_j(\phi f) - (\partial_j \phi) f \in L^p(U)$$

for all $\phi \in C_c^\infty(U)$. By another application of Lemma 1.4, we conclude that also $\partial_j f \in L^p_{loc}(U)$. Hence, $f \in W^{1,p}_{loc}(U)$, as asserted. □

2 The Beurling-Ahlfors Transform

2.1 Motivation

Singular integral operators arise naturally in the study of certain partial differential equations. While this might not directly be clear for the Hilbert transform which we will define in the succeeding subsection, this will be the direct motivation for studying the Beurling-Ahlfors transform and its relation to quasiconformal mappings which are defined as certain solutions to a certain partial differential equation. One can check that the distribution $\mathrm{PV}\, 1/z^2$ defines a tempered distribution in $\mathbf{C}$. Then we can define the Beurling-Ahlfors transform $\mathscr{B}$ as the convolution operator

$$\mathscr{B}\phi(w) := -\frac{1}{\pi}\,\mathrm{PV}\,\frac{1}{z^2} * \phi(w) = -\frac{1}{\pi}\lim_{\varepsilon\downarrow 0}\int_{|w-z|\geq\varepsilon}\frac{\phi(z)}{(w-z)^2}\,\mathrm{d}z$$

and try to deduce properties from this formula. However, the choice of our operator may now seem rather arbitrary. To facilitate a more natural approach we will therefore, in a sense, work in a backwards manner, in particular when compared to the more direct approach we will take for the Hilbert transform in Subsection 2.2.

When working in $\mathbf{C}$ and, especially when working with results from complex analysis, the partial differential operators ∂_x and ∂_y obtained from the natural coordinates of $\mathbf{R}^2$ are not always the natural choice. We define the so-called *Wirtinger derivatives* as the linear differential operators

$$\partial_z := \frac{1}{2}(\partial_x - i\partial_y), \quad \partial_{\overline{z}} := \frac{1}{2}(\partial_x + i\partial_y).$$

In many ways these operators behave like differentiation in $\mathbf{R}$. See also Appendix C and in particular Proposition C.1.

We start with a rather simple example which captures the general idea. Suppose a function $f \in L^2(\mathbf{C})$ satisfies $\partial_{\overline{z}} f \in L^2(\mathbf{C})$. One can then ask if it must also be true that $\partial_z f \in L^2(\mathbf{C})$, and, if yes, if its L^2-norm can be estimated by that of $\partial_{\overline{z}} f$. A way to solve this problem is by finding an operator $\mathscr{L} \in \mathcal{L}(L^2(\mathbf{C}))$ that satisfies

$$\mathscr{L}(\partial_{\overline{z}} f) = \partial_z f.$$

By taking the Fourier transform we obtain

$$\pi i\overline{\zeta}\mathscr{F} f = \mathscr{F}(\partial_z f) = \mathscr{F}\mathscr{L}(\partial_{\overline{z}} f) = \mathscr{F}\mathscr{L}\mathscr{F}^{-1}(\pi i\zeta\mathscr{F} f).$$

This equation is certainly satisfied if we define our operator so that $\mathscr{F}\mathscr{L}\mathscr{F}^{-1}$ is the operator that multiplies a function by $\overline{\zeta}/\zeta$. Since $|\overline{\zeta}/\zeta| = 1$, this operator is an isometry of $L^2(\mathbf{C})$. This positively answers both our questions, with

$$\|\partial_z f\|_2 = \|\mathscr{L}(\partial_{\overline{z}} f)\|_2 = \|\partial_{\overline{z}} f\|_2.$$

Another way of looking at this, is that, since $\partial_x = \partial_z + \partial_{\overline{z}}$, $\partial_y = i(\partial_z - \partial_{\overline{z}})$, we have shown that the domain of the unbounded operator $\partial_{\overline{z}}$ in $L^2(\mathbf{C})$ is precisely the Sobolev space $W^{1,2}(\mathbf{C})$ with the norm $\|f\|_2 + \|\partial_{\overline{z}} f\|_2$, which is equivalent to the usual norm on $W^{1,2}(\mathbf{C})$. One might view this as a form of elliptic regularity of the elliptic partial differential operator $\partial_{\overline{z}}$ in $\mathbf{C}$. Of course, throughout this example one could switch the roles of ∂_z and $\partial_{\overline{z}}$ to deduce an analogous result for the operator ∂_z.

We needed the fact that the Fourier transform is a unitary isomorphism of $L^2(\mathbf{C})$ for this particular argument to work. The operator $\mathscr{L}$ we found in this example is actually equal to the extension to $L^2(\mathbf{C})$ of the Beurling-Ahlfors transform $\mathscr{B}$. As we will show in Subsection 2.3, the Beurling-Ahlfors transform has extensions to $L^p(\mathbf{C})$ for $p \in]1, \infty[$, the proof of which will be facilitated by an analogous result for the Hilbert transform. This result can be used to positively answer the question if for all $p \in]1, \infty[$ there exist constants $c_p \in \mathbf{R}_+$ so that

$$\|\partial_z f\|_p \leq c_p \|\partial_{\overline{z}} f\|_p,$$

for all $f \in L^p(\mathbf{C})$ satisfying $\partial_{\overline{z}} f \in L^p(\mathbf{C})$. By density of $C_c^\infty(\mathbf{C})$ in the Sobolev space $W^{1,p}(\mathbf{C})$, it is actually equivalent to ask if such an inequality holds in the more classical sense where $f \in C_c^\infty(\mathbf{C})$.

Finding the optimal constants c_p however, which are given by $\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))}$, turns out to be quite problematic. For $p \in]1, \infty[$ we write $p^* := \max(p, p')$. Then we do have the following conjecture:

2.1 Conjecture. *For all $p \in]1, \infty[$ we have*

$$\frac{1}{p^*-1}\|\partial_{\overline{z}}\phi\|_p \leq \|\partial_z \phi\|_p \leq (p^*-1)\|\partial_{\overline{z}}\phi\|_p \quad \text{for all } \phi \in C_c^\infty(\mathbf{C}). \tag{2.1}$$

The case $p = 2$ has been shown in our example above. What we do know so far is that if (2.1) is true, then the constants are optimal. Conjecture 2.1 is actually equivalent to Iwaniec's Conjecture:

2.2 Conjecture (Iwaniec)**.** *Let $p \in]1, \infty[$. Then*

$$\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} = p^* - 1 = \begin{cases} p-1 & \text{if } p \in [2, \infty[\\ \dfrac{1}{p-1} & \text{if } p \in]1, 2]. \end{cases}$$

The estimate $\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} \geq p^* - 1$ was already known to Iwaniec and is shown in Proposition 2.45 below. Thus, the conjecture is the upper bound $\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq p^* - 1$. In Section 3 we will go into Iwaniec' motivation for this conjecture. We will use Subsection 2.2 to not only prove L^p-boundedness of the Hilbert transform, but also to specifically determine its operator norms, as this is a deep result of similar type to Iwaniec' Conjecture. We will use Subsection 2.3 to obtain preliminary results, including L^p-boundedness, for the Beurling-Ahlfors transform.

2.2 The Hilbert Transform

The goal of this subsection is to study the Hilbert transform, which shares several properties with the Beurling-Ahlfors transform. Mainly, both operators are integral operators with a comparable singular integration kernel. They are both Fourier multipliers, from which it can easily be seen how they extend to an isometry of $L^2(\mathbf{R})$ and $L^2(\mathbf{C})$ respectively. However, instead of using Fourier analysis we will initially use complex analysis to work out several methods of extending the Hilbert transform to a bounded operator on $L^p(\mathbf{R})$ for $p \in]1, \infty[$ in order to give a fresh presentation on the Hilbert transform, and to see which role it plays in the theory of complex analysis. Additionally, a method due to Calderón and Zygmund called the method of rotations will be used to establish that the Beurling-Ahlfors transform extends to a bounded operator on $L^p(\mathbf{C})$ for $p \in]1,\infty[$ as a direct consequence of the corresponding property of the Hilbert transform.

We define the Hilbert transform $\mathscr{H} : C_c^\infty(\mathbf{R}^n) \to L^0(\mathbf{R}^n)$ as the linear operator

$$\mathscr{H}\phi(x) := \frac{1}{\pi}\,\mathrm{PV}\,\frac{1}{t} * \phi(x) := \frac{1}{\pi}\lim_{\varepsilon\downarrow 0}\int_{|x-t|\geq\varepsilon}\frac{\phi(t)}{x-t}\,\mathrm{d}t.$$

By the Mean Value Theorem we have

$$\left|\frac{\phi(x-t)-\phi(x+t)}{t}\right| = 2\left|\frac{\phi(x-t)-\phi(x+t)}{x-t-(x+t)}\right| \leq 2\|\phi'\|_\infty < \infty. \tag{2.2}$$

Thus, compactness of the support of ϕ justifies writing

$$\begin{aligned}\mathscr{H}\phi(x) &= \lim_{\varepsilon\downarrow 0}\frac{1}{\pi}\left(\int_{-\infty}^{x-\varepsilon}\frac{\phi(t)}{x-t}\,\mathrm{d}t + \int_{x+\varepsilon}^{\infty}\frac{\phi(t)}{x-t}\,\mathrm{d}t\right)\\ &= \lim_{\varepsilon\downarrow 0}\frac{1}{\pi}\int_{\varepsilon}^{\infty}\frac{\phi(x-t)-\phi(x+t)}{t}\,\mathrm{d}t \\ &= \frac{1}{\pi}\int_{\mathbf{R}_+}\frac{\phi(x-t)-\phi(x+t)}{t}\,\mathrm{d}t.\end{aligned} \tag{2.3}$$

Our main theorem of this section is the following:

2.3 Theorem. *Let $p \in]1,\infty[$. Then the Hilbert transform extends to a bounded operator $\mathscr{H} : L^p(\mathbf{R}) \to L^p(\mathbf{R})$ with norm*

$$\|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))} = \begin{cases}\cot\dfrac{\pi}{2p} & \text{if } p \in [2,\infty[\\[2ex] \tan\dfrac{\pi}{2p} & \text{if } p \in]1,2].\end{cases} \tag{2.4}$$

Note that such extensions must be unique, since $C_c^\infty(\mathbf{R})$ is dense in $L^p(\mathbf{R})$ for all $p \in]1,\infty[$, see Theorem A.11.

As it turns out, the Hilbert transform gives a relation between the real and the imaginary part of certain functions which we will use to establish the lower bounds of the operator norms in Theorem 2.3. The main tools we will be using to study this are Cauchy integrals. We denote by $\mathbf{R}_+$ the (strictly) positive real numbers, and by $\mathbf{H}$ the upper half plane of $\mathbf{C}$, i.e., $\mathbf{H} := \{z = x+iy \in \mathbf{C} \mid y \in \mathbf{R}_+\}$. The following theorem states our relation precisely:

2.4 Theorem. *Let $p \in]1,\infty[$ and $g \in L^p(\mathbf{R})$. Suppose $f : \mathbf{H} \to \mathbf{C}$ is a holomorphic function so that*

$$\lim_{y\downarrow 0} f(x+iy) = g(x) \tag{2.5}$$

for a.e $x \in \mathbf{R}$. If there exist $R, c \in \mathbf{R}_+$ so that for $z \in \mathbf{H}$ we have

$$|f(z)| \leq \frac{c}{|z|} \quad \text{if } |z| \geq R, \tag{2.6}$$

then

$$\mathscr{H}(\operatorname{Re} g) = \operatorname{Im} g.$$

The proof of this remarkable theorem will follow naturally through the course of this section and is given below. First we will discuss some details concerning Theorem 2.3.

If for $p, q \in]1, \infty[$ we denote the extensions of the Hilbert transform to $L^p(\mathbf{R})$ and $L^q(\mathbf{R})$ by $\mathscr{H}_p$ and $\mathscr{H}_q$ respectively, then one would wish that whenever $f \in L^p(\mathbf{R}) \cap L^q(\mathbf{R})$, we have $\mathscr{H}_p f = \mathscr{H}_q f$. The following lemma asserts that this must indeed be the case, which means that we can unambiguously denote both operators by $\mathscr{H}$.

2.5 Lemma. *Let $n \in \mathbf{N}$ and $p, q \in]1, \infty[$. Suppose a linear operator $\mathscr{L} : C_c^\infty(\mathbf{R}^n) \to L^0(\mathbf{R}^n)$ has extensions $\mathscr{L}_p \in \mathcal{L}(L^p(\mathbf{R}^n))$ and $\mathscr{L}_q \in \mathcal{L}(L^q(\mathbf{R}^n))$. If $f \in L^p(\mathbf{R}) \cap L^q(\mathbf{R})$, then $\mathscr{L}_p f = \mathscr{L}_q f$.*

Proof. First suppose f has compact support. Let $(\phi_\varepsilon)_{\varepsilon \in \mathbf{R}_+}$ be the standard mollifier, see Definition A.10. Then the sequence $(f_j)_{j \in \mathbf{N}}$ in $C_c^\infty(\mathbf{R}^n)$ defined by $f_j := \phi_{1/j} * f$ has the property that $f_j \to f$ as $j \to \infty$ in both $L^p(\mathbf{R}^n)$ and $L^q(\mathbf{R}^n)$ by Theorem A.7. But then, since $(\mathscr{L}_p f_j)_{j \in \mathbf{N}}$ converges in $L^p(\mathbf{R}^n)$, there is an a.e. convergent subsequence $(\mathscr{L}_p f_{j_k})_{k \in \mathbf{N}}$ with limit $\mathscr{L}_p f$. Moreover, since $(\mathscr{L}_q f_{j_k})_{k \in \mathbf{N}}$ converges in $L^q(\mathbf{R}^n)$, there is an a.e. convergent subsequence $(\mathscr{L}_q f_{j_{k_l}})_{l \in \mathbf{N}}$ with limit $\mathscr{L}_q f$. Thus, taking a.e. limits, we obtain

$$\mathscr{L}_p f = \lim_{l \to \infty} \mathscr{L}_p f_{j_{k_l}} = \lim_{l \to \infty} \mathscr{L}_q f_{j_{k_l}} = \mathscr{L}_q f,$$

since $\mathscr{L}_p$ and $\mathscr{L}_q$ coincide on $C_c^\infty(\mathbf{R}^n)$.

Now suppose $f \in L^p(\mathbf{R}^n) \cap L^q(\mathbf{R}^n)$ is arbitrary. For each $j \in \mathbf{N}$ we denote by χ_j the indicator function of the ball of radius j in $\mathbf{R}^n$. If we set $f_j := \chi_j f$, then it follows from Lebesgue's Dominated Convergence Theorem that $f_j \to f$ as $j \to \infty$ in both $L^p(\mathbf{R}^n)$ and $L^q(\mathbf{R}^n)$. Since f_j has compact support for all $j \in \mathbf{N}$, we may conclude from our previous result, and by using an analogous subsubsequence argument, that $\mathscr{L}_p f = \mathscr{L}_q f$. This proves the desired result. □

The following proposition uses some basic functional analysis to prepare us for the proof of Theorem 2.3.

2.6 Proposition. *Let $p \in]1, \infty[$ and let $p' = p/(p-1) \in]1, \infty[$ denote the conjugate exponent of p. Suppose $\mathscr{H}$ extends to an operator $\mathscr{H} \in \mathcal{L}(L^p(\mathbf{R}))$. Then $\mathscr{H}$ also extends to an operator in $\mathcal{L}(L^{p'}(\mathbf{R}))$, where the extension is given through the dual operator by $-\mathscr{H}^*$. Moreover, if we have established (2.4) for p, then it also holds for p'.*

Proof. Let $\varepsilon \in \mathbf{R}_+$ and $\phi, \psi \in C_c^\infty(\mathbf{R})$. Then,

$$\begin{aligned}
\int_\varepsilon^\infty \int_{\mathbf{R}} \frac{\phi(x-t) - \phi(x+t)}{t} \psi(x) \,\mathrm{d}x \,\mathrm{d}t &= \int_\varepsilon^\infty \int_{\mathbf{R}} \frac{\phi(x-t)\psi(x)}{t} \,\mathrm{d}x \,\mathrm{d}t - \int_\varepsilon^\infty \int_{\mathbf{R}} \frac{\phi(x+t)\psi(x)}{t} \,\mathrm{d}x \,\mathrm{d}t \\
&= \int_\varepsilon^\infty \int_{\mathbf{R}} \frac{\phi(y)\psi(y+t)}{t} \,\mathrm{d}y \,\mathrm{d}t - \int_\varepsilon^\infty \int_{\mathbf{R}} \frac{\phi(y)\psi(y-t)}{t} \,\mathrm{d}y \,\mathrm{d}t \\
&= -\int_\varepsilon^\infty \int_{\mathbf{R}} \frac{\psi(y-t) - \psi(y+t)}{t} \phi(y) \,\mathrm{d}y \,\mathrm{d}t.
\end{aligned}$$

As both ϕ and ψ have compact support we are justified in letting $\varepsilon \downarrow 0$ to conclude from (2.3) and Fubini's Theorem that

$$\langle \phi, \mathscr{H}^* \psi \rangle = \langle \mathscr{H} \phi, \psi \rangle = \langle \phi, -\mathscr{H} \psi \rangle. \tag{2.7}$$

Since, by Lemma 1.1, we have

$$\sup_{\substack{\phi,\psi\in C_c^\infty(\mathbf{R})\\ \|\phi\|_p=\|\psi\|_{p'}=1}} |\langle \phi, -\mathscr{H}\psi\rangle| = \sup_{\substack{\phi,\psi\in C_c^\infty(\mathbf{R})\\ \|\phi\|_p=\|\psi\|_{p'}=1}} |\langle \phi, \mathscr{H}^*\psi\rangle| = \|\mathscr{H}^*\|_{\mathcal{L}(L^{p'}(\mathbf{R}))} < \infty,$$

it now follows that $\mathscr{H}$ has an extension in $\mathcal{L}(L^{p'}(\mathbf{R}))$ and, by (2.7), $-\mathscr{H}^*|_{C_c^\infty(\mathbf{R})} = \mathscr{H}$.

Since

$$\| -\mathscr{H}^*\|_{\mathcal{L}(L^{p'}(\mathbf{R}))} = \|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))},$$

and by noting that for conjugate exponents $q, q' \in]1,\infty[$ one has

$$\cot\frac{\pi}{2q} = \cot\left(\frac{\pi}{2}\left(1-\frac{1}{q'}\right)\right) = \tan\frac{\pi}{2q'},$$

the assertion follows.

□

From this proposition we may conclude that it suffices to establish Theorem 2.3 for the cases where $p \in]1,2]$ or $p \in [2,\infty[$, since the map $p \mapsto p/(p-1)$ gives a bijection from $]1,2]$ to $[2,\infty[$ and vice versa.

2.7 Definition. Let $f \in L^0(\mathbf{R})$ so that $t \mapsto f(t)/(t-z)$ is in $L^1(\mathbf{R})$ for all $z \in \mathbf{H}$. Then the function $Cf : \mathbf{H} \to \mathbf{C}$ defined by

$$Cf(z) := \frac{1}{\pi i}\int_{\mathbf{R}} \frac{f(t)}{t-z}\,\mathrm{d}t$$

is called the *Cauchy integral* of f. ♢

2.8 Lemma. *Let $p \in [1,\infty[$ and $f \in L^p(\mathbf{R})$. Then Cf is well-defined. Moreover, there is a continuous map $c_p : \mathbf{R}_+ \to \mathbf{R}_+$, depending only on $p \in [1,\infty[$, so that*

$$|Cf(z)| \le c_p(y)\|f\|_p$$

for all $z = x + iy \in \mathbf{H}$.

Proof. Let $f \in L^p(\mathbf{R})$ for $p \in [1,\infty[$ and $z = x+iy \in \mathbf{H}$. Then, noting that $|t-z|^2 = (t-x)^2 + y^2$ majorizes both y^2 and $(t-x)^2$ for all $t \in \mathbf{R}$, we find, if $p > 1$,

$$\begin{aligned}
\int_{\mathbf{R}} \frac{|f(t)|}{|t-z|}\,\mathrm{d}t &\le \int_{|x-t|\le 1} \frac{|f(t)|}{y}\,\mathrm{d}t + \int_{|x-t|\ge 1} \frac{|f(t)|}{|t-x|}\,\mathrm{d}t\\
&\le \frac{2^{\frac{1}{p'}}\|f\|_p}{y} + \|f\|_p \int_{|t|\ge 1} |t|^{-p'}\,\mathrm{d}t\\
&= \left(\frac{2^{\frac{1}{p'}}}{y} + \frac{2}{p'-1}\right)\|f\|_p,
\end{aligned}$$

by Hölders inequality, and

$$\int_{\mathbf{R}} \frac{|f(t)|}{|t-z|}\,\mathrm{d}t \le \frac{\|f\|_1}{y}$$

if $p=1$. This means that Cf is well-defined. By setting

$$c_p(y) := \frac{2^{1-\frac{1}{p}}}{\pi y} + \frac{2}{\pi}(p-1),$$

the assertion follows. □

By the above lemma we may use Lebesgue's Dominated Convergence Theorem to take limits under the integral sign to conclude that Cf is continuous. As a matter of fact, we have the following:

2.9 Lemma. *Let $p \in [1,\infty[$ and $f \in L^p(\mathbf{R})$. Then Cf is holomorphic in $\mathbf{H}$.*

Proof. For each $t \in \mathbf{R}$ one notes that the map $z \mapsto f(t)/(t-z)$ is holomorphic in $\mathbf{H}$. Then it follows from Cauchy's Integral Theorem that for any closed contour $\Gamma \subseteq \mathbf{H}$ and all $t \in \mathbf{R}$ we have

$$\oint_\Gamma \frac{f(t)}{t-z}\,\mathrm{d}z = 0.$$

By Lemma 2.8 we are justified in applying Fubini's Theorem to find that

$$\oint_\Gamma Cf(z)\,\mathrm{d}z = \frac{1}{\pi i}\int_{\mathbf{R}}\oint_\Gamma \frac{f(t)}{t-z}\,\mathrm{d}z\,\mathrm{d}t = 0$$

for any closed contour $\Gamma \subseteq \mathbf{H}$. The result now follows from Morera's Theorem, see Theorem C.23. □

2.10 Definition. We define the *Poisson kernel* $(P_y)_{y\in\mathbf{R}_+}$ and the *associated Poisson kernel* $(Q_y)_{y\in\mathbf{R}_+}$ by

$$P_y(x) := P(x,y) = \operatorname{Re} -\frac{1}{\pi i z} = \frac{1}{\pi}\frac{y}{x^2+y^2},$$
$$Q_y(x) := Q(x,y) = \operatorname{Im} -\frac{1}{\pi i z} = \frac{1}{\pi}\frac{x}{x^2+y^2},$$

where $z = x+iy \in \mathbf{H}$. ♢

As P and Q are respectively the real and the imaginary part of a holomorphic function, they are associated harmonic functions in $\mathbf{H}$. Now let $u \in C_c^\infty(\mathbf{R})$ be a fixed real-valued function. Then we can write

$$Cu(z) = (P_y * u)(x) + i(Q_y * u)(x) \quad (z \in \mathbf{H}), \tag{2.8}$$

where the real and imaginary parts are given by

$$(P_y * u)(x) = \frac{1}{\pi}\int_{\mathbf{R}} u(t)\frac{y}{(x-t)^2+y^2}\,\mathrm{d}t,$$

$$(Q_y * u)(x) = \frac{1}{\pi}\int_{\mathbf{R}} u(t)\frac{x-t}{(x-t)^2+y^2}\,\mathrm{d}t \quad (z \in \mathbf{H}).$$

The following lemma shows that the function $u + i\mathscr{H}u$ gives the boundary values of Cu along the real axis.

2.11 Lemma. *Let $p \in [1,\infty[$ and $f \in L^p(\mathbf{R})$. Then*

$$\lim_{y\downarrow 0} P_y * f = f, \tag{2.9}$$

with limit in $L^p(\mathbf{R})$. For any $f \in C_0(\mathbf{R})$ the limit (2.9) holds in $L^\infty(\mathbf{R})$. Furthermore, for any $\phi \in C_c^\infty(\mathbf{R})$ we have

$$\lim_{y\downarrow 0} Q_y * \phi = \mathscr{H}\phi,$$

where the limit is in $L^\infty(\mathbf{R})$.

Proof. We set $P := P_1$. Then

$$\int_{\mathbf{R}^n} P(x)\,\mathrm{d}x = \frac{1}{\pi}\int_{\mathbf{R}} \frac{1}{x^1+1}\,\mathrm{d}x = \frac{1}{\pi}\left[\arctan(x)\right]_{-\infty}^{\infty} = 1.$$

Since $P_y(x) = y^{-1}P(x/y)$ for all $x \in \mathbf{R}$, $y \in \mathbf{R}_+$, it follows from Proposition A.8 that the family $(P_y)_{y\in\mathbf{R}_+}$ is an approximate identity. The assertions about P_y then follow from Theorem A.7.

Fix $x \in \mathbf{R}$. For the assertion about Q_y, we will rewrite $Q_y * \phi$ as

$$(Q_y * \phi)(x) = \frac{1}{\pi}\int_{\mathbf{R}} \phi(x-t)\frac{t}{t^2+y^2}\,\mathrm{d}t = \frac{1}{\pi}\int_{\mathbf{R}_+} (\phi(x-t)-\phi(x+t))\frac{t}{t^2+y^2}\,\mathrm{d}t. \tag{2.10}$$

By combining (2.10) and (2.3) we obtain, by (2.2) and by noting that $-\partial_t \arctan y/t = y/(t^2+y^2)$,

$$\begin{aligned}
|(Q_y * \phi)(x) - \mathscr{H}\phi(x)| &\le \frac{2}{\pi}\|\phi'\|_\infty \int_{\mathbf{R}_+} \left|\frac{t^2}{t^2+y^2} - 1\right| \mathrm{d}t \\
&= \frac{2y}{\pi}\|\phi'\|_\infty \int_{\mathbf{R}_+} \frac{y}{t^2+y^2}\,\mathrm{d}t \\
&= y\|\phi'\|_\infty.
\end{aligned}$$

Hence, since x was arbitrary,

$$\|Q_y * \phi - \mathscr{H}\phi\|_{L^\infty(\mathbf{R})} \le y\|\phi'\|_\infty \to 0 \quad \text{as } y \downarrow 0,$$

proving the assertion. □

By (2.8) and Lemma 2.11, we may conclude that for any $u \in C_c^\infty(\mathbf{R})$ we have

$$\lim_{y\downarrow 0} Cu(x+iy) = u(x) + i(\mathscr{H}u)(x), \tag{2.11}$$

uniformly in $x \in \mathbf{R}$. This limit will allow us to use complex contour integration to establish L^p bounds for the Hilbert transform, which establishes the first assertion of Theorem 2.3.

2.12 Proposition. *Let $p \in]1,\infty[$. Then the Hilbert transform extends to a bounded operator $\mathscr{H} : L^p(\mathbf{R}) \to L^p(\mathbf{R})$. Furthermore, the extension to $L^2(\mathbf{R})$ is isometric.*

An important tool we will use for the proof is the Riesz-Thorin Interpolation Theorem. We will use the following version of this theorem:

2.13 Theorem (Riesz-Thorin)**.** *Let $p, q \in [1,\infty]$, $p \le q$. Suppose a non-zero linear operator $\mathscr{L} : C_c^\infty(\mathbf{R}^n) \to L^0(\mathbf{R}^n)$ has extensions $\mathscr{L}_p \in \mathcal{L}(L^p(\mathbf{R}^n))$ and $\mathscr{L}_q \in \mathcal{L}(L^q(\mathbf{R}^n))$. Then, for each $r \in [p, q]$ there is an extension $\mathscr{L}_r \in \mathcal{L}(L^r(\mathbf{R}^n))$ of $\mathscr{L}$. Moreover, the function*

$$\left[\frac{1}{q}, \frac{1}{p}\right] \to \mathbf{R}, \quad t \mapsto \log \|\mathscr{L}_{t^{-1}}\|_{\mathcal{L}(L^{t^{-1}}(\mathbf{R}^n))}$$

is convex.

This version of the theorem is a consequence of the general result as can, for example, be found in [Gr, Theorem 1.3.4], together with Lemma 2.5. With the notation as in the theorem, the convexity condition is usually stated as

$$\|\mathscr{L}_r\|_{\mathcal{L}(L^r(\mathbf{R}^n))} \le \|\mathscr{L}_p\|^{1-t}_{\mathcal{L}(L^p(\mathbf{R}^n))} \|\mathscr{L}_q\|^{t}_{\mathcal{L}(L^q(\mathbf{R}^n))}$$

where $t \in [0, 1]$ and r is given through

$$\frac{1}{r} = \frac{(1-t)}{p} + \frac{t}{q}.$$

As is common for arguments involving complex contour integrals, we will need to establish the behavior of Cu at infinity.

2.14 Lemma. *Let $u \in C_c^\infty(\mathbf{R})$. There exist $R, c \in \mathbf{R}_+$ so that for all $z \in \mathbf{H}$ with $|z| \ge R$ we have*

$$|Cu(z)| \le \frac{c}{|z|}.$$

Proof. We note that

$$\left|\frac{z}{t-z}\right| = \left|\frac{1}{\frac{t}{z} - 1}\right| \to 1 \quad \text{as } |z| \to \infty,$$

uniformly for $t \in \operatorname{supp} u$. Hence,

$$\lim_{|z|\to\infty} |zCu(z)| = \lim_{|z|\to\infty} \left|\frac{1}{\pi i}\int_{\mathbf{R}} u(t)\frac{z}{t-z}\,\mathrm{d}t\right| = \frac{1}{\pi}\left|\int_{\mathbf{R}} u(t)\,\mathrm{d}t\right| =: c'.$$

Setting $c := c' + 1 \in \mathbf{R}_+$, the existence of $R \in \mathbf{R}_+$ is a consequence of the definition of the limit. □

2.15 Lemma. *Let $u \in C_c^\infty(\mathbf{R})$ and let $R, c \in \mathbf{R}_+$ be as in Lemma 2.14. Then for all $x \in \mathbf{R}$ with $|x| \ge R$ we have*

$$|u(x) + i(\mathscr{H}u)(x)| \le \frac{c}{|x|}.$$

Proof. Suppose $x \in \mathbf{R}$ satisfies $|x| \ge R$ and let $\varepsilon \in \mathbf{R}_+$. Then, by (2.11), we can choose $y \in \mathbf{R}_+$ small enough so that

$$|u(x) + i(\mathscr{H}u)(x) - Cu(x+iy)| < \varepsilon. \tag{2.12}$$

Moreover, we have $|x+iy| > |x| \ge R$. Hence,

$$|Cu(x+iy)| \le \frac{c}{|x+iy|} < \frac{c}{|x|}. \tag{2.13}$$

Thus, by (2.12) and (2.13),

$$|u(x) + i(\mathscr{H}u)(x)| \leq |u(x) + i(\mathscr{H}u)(x) - Cu(x+iy)| + |Cu(x+iy)| < \varepsilon + \frac{c}{|x|}.$$

The assertion follows by letting $\varepsilon \downarrow 0$. □

Proof of Proposition 2.12. By Proposition 2.6 it suffices to check the cases where $p \in [2, \infty[$. By the Riesz-Thorin Interpolation Theorem it is then sufficient to find an increasing sequence $(p_k)_{k\in\mathbf{N}}$ with $p_1 = 2$ and $p_k \to \infty$ as $k \to \infty$ for which the extensions exist. For our proof we will consider the cases where $p = 2k$, for $k \in \mathbf{N}$.

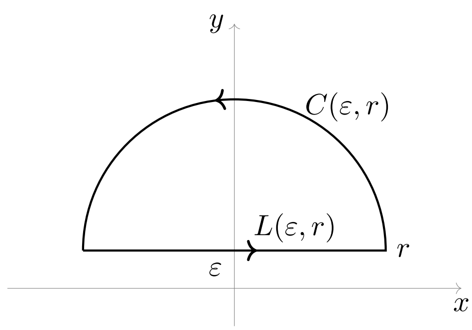
Fix $k \in \mathbf{N}$ and let $\varepsilon, r \in \mathbf{R}_+$. Denote by $\Gamma(\varepsilon, r)$ the closed contour in $\mathbf{H}$ given by the union of line segment

$$L(\varepsilon, r) := \{x + i\varepsilon \in \mathbf{H} \,|\, x \in [-r, r]\}$$

and the semicircle

$$C(\varepsilon, r) := \left\{ re^{2\pi i t} + i\varepsilon \,\middle|\, t \in \left[0, \frac{1}{2}\right] \right\},$$

oriented as in the figure below.



Let $u \in C_c^\infty(\mathbf{R})$ be real-valued. Then $(Cu)^{2k}$ is analytic in $\mathbf{H}$, so it follows from Cauchy's Theorem that

$$\oint_{\Gamma(\varepsilon,r)} Cu(z)^{2k} \, \mathrm{d}z = 0.$$

We will show that

$$\int_{\mathbf{R}} (u(x) + i\mathscr{H}u(x))^{2k} \, \mathrm{d}x = \lim_{r\to\infty} \lim_{\varepsilon\downarrow 0} \oint_{\Gamma(\varepsilon,r)} Cu(z)^{2k} \, \mathrm{d}z = 0. \tag{2.14}$$

First we will have to show that $(u(x) + i\mathscr{H}u(x))^{2k} \in L^1(\mathbf{R})$. Let $R, c \in \mathbf{R}_+$ be as in Lemma 2.14. Then it suffices to note that, by Lemma 2.15, we have

$$\begin{aligned}\int_{\mathbf{R}} |u(x) + i\mathscr{H}u(x)|^{2k} \, \mathrm{d}x &\leq \int_{|x|\leq R} |u(x) + i\mathscr{H}u(x)|^{2k} \, \mathrm{d}x + \int_{|x|>R} \frac{c^{2k}}{|x|^{2k}} \\ &= \int_{|x|\leq R} |u(x) + i\mathscr{H}u(x)|^{2k} \, \mathrm{d}x + \frac{2c^{2k}}{2k-1} \frac{1}{R^{2k-1}} < \infty.\end{aligned}$$

Next, letting $r \geq R$, we note that $R \leq r = |re^{2\pi it}| \leq |re^{2\pi it} + i\varepsilon|$ for all $\varepsilon \in \mathbf{R}_+$ and $t \in [0, 1/2]$. This allows us to estimate

$$|Cu(re^{2\pi it} + i\varepsilon)|^{2k}|re^{2\pi it}| \leq \frac{rc^{2k}}{|re^{2\pi it} + i\varepsilon|^{2k}} \leq \frac{c^{2k}}{r^{2k-1}},$$

which, on account of Lebesgue's Dominated Convergence Theorem, allows us to conclude both the estimate and the existence of the limit in

$$\lim_{\varepsilon \downarrow 0} \left| \int_{C(\varepsilon,r)} Cu(z)^{2k} \, \mathrm{d}z \right| \leq 2\pi \lim_{\varepsilon \downarrow 0} \int_0^{\frac{1}{2}} |Cu(re^{2\pi it} + i\varepsilon)|^{2k} |re^{2\pi it}| \, \mathrm{d}t \leq \frac{\pi c^{2k}}{r^{2k-1}}.$$

It follows that

$$\lim_{r \to \infty} \lim_{\varepsilon \downarrow 0} \int_{C(\varepsilon,r)} Cu(z)^{2k} \, \mathrm{d}z = 0. \tag{2.15}$$

Next, we note that for any $r \in \mathbf{R}_+$ we have

$$\lim_{\varepsilon \downarrow 0} \int_{L(\varepsilon,r)} Cu(z)^{2k} \, \mathrm{d}z = \lim_{\varepsilon \downarrow 0} \int_{|x| \leq r} Cu(x + i\varepsilon)^{2k} \, \mathrm{d}x = \int_{|x| \leq r} (u(x) + i\mathscr{H}u(x))^{2k} \, \mathrm{d}x,$$

where the interchange of the order of the limit and integration is justified by the uniform convergence in (2.11). Then, since $(u + i\mathscr{H}u)^{2k} \in L^1(\mathbf{R})$, we may conclude that

$$\lim_{r \to \infty} \lim_{\varepsilon \downarrow 0} \int_{L(\varepsilon,r)} Cu(z)^{2k} \, \mathrm{d}z = \int_{\mathbf{R}} (u(x) + i\mathscr{H}u(x))^{2k} \, \mathrm{d}x. \tag{2.16}$$

Combining (2.15) and (2.16) yields (2.14).

Now set $k = 1$. Then, by taking the real part of (2.14), we find that

$$\int_{\mathbf{R}} (u(x)^2 - \mathscr{H}u(x)^2) \, \mathrm{d}x = 0$$

This proves that for any real-valued $u \in C_c^\infty(\mathbf{R})$ we have $\|\mathscr{H}u\|_2 = \|u\|_2$. Now let $\phi \in C_c^\infty(\mathbf{R})$ be arbitrary. Then, since $\mathscr{H}$ maps real-valued functions to real-valued functions, we have

$$\begin{aligned}\|\mathscr{H}\phi\|_2^2 &= \int_{\mathbf{R}} \mathscr{H}(\operatorname{Re}\phi)(x)^2 \, \mathrm{d}x + \int_{\mathbf{R}} \mathscr{H}(\operatorname{Im}\phi)(x)^2 \, \mathrm{d}x \\ &= \int_{\mathbf{R}} ((\operatorname{Re}\phi(x))^2 + (\operatorname{Im}\phi(x))^2) \, \mathrm{d}x = \|\phi\|_2^2.\end{aligned}$$

Then, by Lemma 1.1, the Hilbert transform extends to $L^2(\mathbf{R})$. Moreover, we may conclude that this extension is isometric.

Now suppose $k > 1$. By the Binomial Theorem, taking the real part of (2.14) yields

$$\int_{\mathbf{R}} \sum_{j=0}^{k} \binom{2k}{2j} (-1)^j u(x)^{2k-2j} \mathscr{H}u(x)^{2j} \, \mathrm{d}x = 0.$$

Hence,

$$\int_{\mathbf{R}} |\mathscr{H}u(x)|^{2k}\,\mathrm{d}x \le \int_{\mathbf{R}} |u(x)|^{2k}\,\mathrm{d}x + \int_{\mathbf{R}} \sum_{j=1}^{k-1} \binom{2k}{2j} |u(x)|^{2k-2j} |\mathscr{H}u(x)|^{2j}\,\mathrm{d}x. \tag{2.17}$$

Now let $j \in \{1, \dots, m-1\}$ and $\varepsilon \in \mathbf{R}_+$ a number smaller than $\sum_{j=1}^{k-1} \binom{2k}{2j}$. If we set $p = k/(k-j)$, then $p' = p/(p-1) = k/j$. Young's inequality asserts that for all $a, b \in \mathbf{R}_{\ge 0}$ we have

$$ab \le \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

By applying this to $a = (\varepsilon p')^{-1/p'} |u(x)|^{2k-2j}$, $b = (\varepsilon p')^{1/p'} |\mathscr{H}u(x)|^{2j}$ for $x \in \mathbf{R}$, we find that there is some constant $c(\varepsilon, j) \in \mathbf{R}_+$ so that

$$|u(x)|^{2k-2j} |\mathscr{H}u(x)|^{2j} \le c(\varepsilon, j) |u(x)|^{2k} + \varepsilon |\mathscr{H}u(x)|^{2k}.$$

Thus, by (2.17), we have

$$\int_{\mathbf{R}} |\mathscr{H}u(x)|^{2k}\,\mathrm{d}x \le \left(1 + \sum_{j=1}^{k-1} \binom{2k}{2j} c(\varepsilon, j)\right) \int_{\mathbf{R}} |u(x)|^{2k}\,\mathrm{d}x + \varepsilon \sum_{j=1}^{k-1} \binom{2k}{2j} \int_{\mathbf{R}} |\mathscr{H}u(x)|^{2k}\,\mathrm{d}x.$$

This proves that there is some $c \in \mathbf{R}_+$ so that $\|\mathscr{H}u\|_{2k}^{2k} \le c\|u\|_{2k}^{2k}$.

Now let $\phi \in C_c^\infty(\mathbf{R})$ be arbitrary. Another similar application of the Binomial Theorem and Young's inequality shows that there is some $c' \in \mathbf{R}_+$ so that for all $a, b \in \mathbf{R}_{\ge 0}$ we have $(a+b)^k \le c'(a^k + b^k)$. Hence, we have

$$\begin{aligned}
\|\mathscr{H}\phi\|_{2k}^{2k} &= \int_{\mathbf{R}} (\mathscr{H}(\operatorname{Re}\phi)(x)^2 + \mathscr{H}(\operatorname{Im}\phi)(x)^2)^k\,\mathrm{d}x \\
&\le c'c \int_{\mathbf{R}} ((\operatorname{Re}\phi)(x)^{2k} + (\operatorname{Im}\phi)(x)^{2k})\,\mathrm{d}x \\
&\le c'c \int_{\mathbf{R}} ((\operatorname{Re}\phi)(x)^2 + (\operatorname{Im}\phi)(x)^2)^k\,\mathrm{d}x = cc'\|\phi\|_{2k}^{2k}.
\end{aligned}$$

The conclusion now follows from Lemma 1.1. □

To obtain a more precise bound on the norm of the Hilbert transform, we will prove a result that will allow us to compute the Hilbert transform of certain functions.

2.16 Lemma. *Suppose $f : \mathbf{H} \to \mathbf{C}$ is a holomorphic function so that*

$$g(x) := \lim_{y \downarrow 0} f(x + iy)$$

exists for a.e. $x \in \mathbf{R}$. If there exist $R, c \in \mathbf{R}_+$ so that for $z \in \mathbf{H}$ we have

$$|f(z)| \le \frac{c}{|z|} \quad \text{if } |z| \ge R, \tag{2.18}$$

and

$$\int_{|x|\le R} \frac{|g(x)|}{|x-w|}\,\mathrm{d}x < \infty \quad \textit{for all } w \in \mathbf{H}, \tag{2.19}$$

then $x \mapsto g(x)/(x-w)$ *is in* $L^1(\mathbf{R})$ *for all* $w \in \mathbf{H}$ *and*

$$f(z) = \frac{1}{2} Cg(z) \tag{2.20}$$

for all $z \in \mathbf{H}$.

Note that (2.19) is automatically satisfied if g is continuous, or, by Lemma 2.8, if $g \in L^p(\mathbf{R})$ for some $p \in [1, \infty[$.

Proof. For $\varepsilon, r \in \mathbf{R}_+$, let $\Gamma(\varepsilon, r) = C(\varepsilon, r) \cup L(\varepsilon, r)$ be the closed contour from the proof of Proposition 2.12. By Cauchy's Integral Formula we find that for all w in the interior of $\Gamma(\varepsilon, r)$ we have

$$f(w) = \frac{1}{2\pi i} \oint_{\Gamma(\varepsilon,r)} \frac{f(z)}{z-w}\,\mathrm{d}z. \tag{2.21}$$

The strategy will be to justify letting $\varepsilon \downarrow 0$ and $r \to \infty$ to obtain (2.20).

Fix $w \in \mathbf{H}$. Choosing $\varepsilon \in \mathbf{R}_+$ small enough and $r \ge R$ large enough, ensures that w lies in the interior of $\Gamma(\varepsilon, r)$. Note that $R \le r = |re^{2\pi it}| \le |re^{2\pi it} + i\varepsilon|$ for all $\varepsilon \in \mathbf{R}_+$ and $t \in [0, 1/2]$. Any $z \in C(\varepsilon, r)$ is of the form $z = re^{2\pi it} + i\varepsilon$ for some $t \in [0, 1/2]$, while any w in the interior of $\Gamma(\varepsilon, r)$ is of the form $\rho e^{2\pi it'} + i\varepsilon$ for some $\rho \in]0, r[$ and $t' \in]0, 1/2[$. This implies, by the reverse triangle inequality, that

$$|z - w| = |re^{2\pi it} - \rho e^{2\pi it'}| \ge r - \rho > 0$$

for all $z \in C(\varepsilon, r)$. Hence, by (2.18),

$$\frac{|f(z)|}{|z-w|} \le \frac{c}{r} \frac{1}{r-\rho}$$

for all $z \in C(\varepsilon, r)$. Then, as a consequence of Lebesgue's Dominated Convergence Theorem, this allows us to conclude both the estimate and the existence of the limit in

$$\lim_{\varepsilon \downarrow 0} \left| \int_{C(\varepsilon,r)} \frac{f(z)}{z-w}\,\mathrm{d}z \right| \le \frac{\pi c}{r-\rho}.$$

Hence,

$$\lim_{r\to\infty} \lim_{\varepsilon\downarrow 0} \int_{C(\varepsilon,r)} \frac{f(t)}{z-w}\,\mathrm{d}z = 0. \tag{2.22}$$

Next, we make the observation that we can assume that $R > |w|$. Indeed, if it were the case that $R \le |w|$, then one notes that by (2.18) the function $x \mapsto g(x)/(x-w)$ is bounded on the annulus $R \le |x| \le |w| + 1$, hence the condition (2.19) will still hold if we replace R by $|w| + 1$. As this change will also not affect (2.18), we can indeed assume that $R > |w|$.

If we assume that $r \ge R$, then, for all $x \in \mathbf{R}$ with $|x| > r$, we have

$$|x - w| \ge |x| - |w| = (|x|^{\frac{1}{2}} + |w|^{\frac{1}{2}})(|x|^{\frac{1}{2}} - |w|^{\frac{1}{2}}) > |x|^{\frac{1}{2}}(r^{\frac{1}{2}} - |w|^{\frac{1}{2}}) > 0$$

Moreover, we have

$$|g(x)| \leq \frac{c}{|x|} \quad \text{for a.e. } x \in \mathbf{R} \text{ satisfying } |x| \geq R.$$

Hence, one shows that $x \mapsto g(x)/(x-w)$ is in $L^1(\mathbf{R})$ by the estimate

$$\begin{aligned}\int_{\mathbf{R}} \frac{|g(x)|}{|x-w|}\,\mathrm{d}x &\leq \int_{|x|\leq R} \frac{|g(x)|}{|x-w|}\,\mathrm{d}x + \int_{|x|>R} \frac{c}{|x|^{\frac{3}{2}}(R^{\frac{1}{2}} - |w|^{\frac{1}{2}})}\,\mathrm{d}x \\ &= \int_{|x|\leq R} \frac{|g(x)|}{|x-w|}\,\mathrm{d}x + \frac{4c}{R^{\frac{1}{2}}(R^{\frac{1}{2}} - |w|^{\frac{1}{2}})} < \infty,\end{aligned}$$

where we used (2.19). Then, by a similar computation, we can conclude that

$$\left| \int_{|x|>r} \frac{g(x)}{x-w}\,\mathrm{d}x \right| \leq \frac{4c}{r^{\frac{1}{2}}(r^{\frac{1}{2}} - |w|^{\frac{1}{2}})} \to 0 \quad \text{as } R \leq r \to \infty.$$

Also noting that for $r \in \mathbf{R}_+$ we have

$$\lim_{\varepsilon \downarrow 0} \int_{L(\varepsilon,r)} \frac{f(z)}{z-w}\,\mathrm{d}z = \lim_{\varepsilon \downarrow 0} \int_{|x|\leq r} \frac{f(x+i\varepsilon)}{x+i\varepsilon-w}\,\mathrm{d}x = \int_{|x|\leq r} \frac{g(x)}{x-w}\,\mathrm{d}x$$

by Lebesgue's Dominated Convergence Theorem, we may conclude that

$$\lim_{r\to\infty} \lim_{\varepsilon \downarrow 0} \int_{L(\varepsilon,r)} \frac{f(z)}{z-w}\,\mathrm{d}z = \lim_{r\to\infty} \int_{|x|\leq r} \frac{g(x)}{x-w}\,\mathrm{d}x = \int_{\mathbf{R}} \frac{g(x)}{x-w}\,\mathrm{d}x. \tag{2.23}$$

Combining (2.21), (2.22), and (2.23) yields

$$f(w) = \frac{1}{2\pi i} \int_{\mathbf{R}} \frac{g(x)}{x-w}\,\mathrm{d}x = \frac{1}{2} Cg(w).$$

The assertion follows. □

As an application of this result, we will prove a remarkable identity known as Cotlar's Identity.

2.17 Proposition (Cotlar's Identity)**.** *For any $\phi \in C_c^\infty(\mathbf{R})$ we have*

$$(\mathscr{H}\phi)^2 = \phi^2 + 2\mathscr{H}(\phi\mathscr{H}\phi). \tag{2.24}$$

Note that $\mathscr{H}\phi \in C^\infty(\mathbf{R})$ by differentiation under the integral sign. Hence, $\phi\mathscr{H}\phi \in C_c^\infty(\mathbf{R})$, meaning that the expression $\mathscr{H}(\phi\mathscr{H}\phi)$ in (2.24) makes sense.

Proof. Let $u \in C_c^\infty(\mathbf{R})$ be real-valued. Then the function $f := i(Cu)^2$ is analytic in $\mathbf{H}$. For $x \in \mathbf{R}$, it follows from (2.11) that this function takes boundary values

$$g(x) := \lim_{y\downarrow 0} f(x+iy) = i(u(x) + i\mathscr{H}u(x))^2 = -2u(x)\mathscr{H}u(x) + i(u(x)^2 - (\mathscr{H}u(x))^2).$$

It now suffices to show that $\mathscr{H}(\operatorname{Re} g) = \operatorname{Im} g$, which yields (2.24) for the real-valued case.

Since Cu satisfies (2.18) by Lemma 2.14, so does f. Hence, by Lemma 2.16, we have

$$\begin{aligned} 2f(x+iy) &= (P_y * g)(x) + i(Q_y * g)(x) \\ &= (P_y * \operatorname{Re} g)(x) - (Q_y * \operatorname{Im} g)(x) + i((P_y * \operatorname{Im} g)(x) + (Q_y * \operatorname{Re} g)(x)). \end{aligned} \tag{2.25}$$

Since $\operatorname{Re} g \in C_c^\infty(\mathbf{R})$, we find, by Lemma 2.11, that $(Q_y * \operatorname{Re} g)(x) \to \mathscr{H}(\operatorname{Re} g)(x)$ as $y \downarrow 0$ for all $x \in \mathbf{R}$. Moreover, since $u \in C_c^\infty(\mathbf{R}) \subseteq L^2(\mathbf{R})$, we also have $\mathscr{H}u \in L^2(\mathbf{R})$ by Proposition 2.12. Hence, $u \pm \mathscr{H}u \in L^2(\mathbf{R})$ so that $\operatorname{Im} g = (u + \mathscr{H}u)(u - \mathscr{H}u) \in L^1(\mathbf{R})$. Thus, it follows from Lemma 2.11 that $P_y * \operatorname{Im} g \to \operatorname{Im} g$ in $L^1(\mathbf{R})$. This implies that the sequence $(P_{1/j} * \operatorname{Im} g)_{j \in \mathbf{N}}$ has an a.e. convergent subsequence with limit $\operatorname{Im} g$. By taking the imaginary part of (2.25) we may pass to this a.e. convergent subsequence to conclude that

$$2 \operatorname{Im} g = \operatorname{Im} g + \mathscr{H}(\operatorname{Re} g)$$

and hence

$$\mathscr{H}(\operatorname{Re} g) = \operatorname{Im} g,$$

as desired. This proves (2.24) for real-valued $\phi = u \in C_c^\infty(\mathbf{R})$.

To obtain the complex valued case, we let $\phi \in C_c^\infty(\mathbf{R})$ be arbitrary. Now fix $x \in \mathbf{R}$ and define the (real) symmetric bilinear forms α, β by

$$\alpha(u,v) = \mathscr{H}u(x)\mathscr{H}v(x), \quad \beta(u,v) = u(x)v(x) + \mathscr{H}(u\mathscr{H}v)(x) + \mathscr{H}(v\mathscr{H}u)(x)$$

for real-valued $u, v \in C_c^\infty(\mathbf{R})$. We have shown that $\alpha(w,w) = \beta(w,w)$ for all real-valued $w \in C_c^\infty(\mathbf{R})$. But then, by the (real) polarization identity, we obtain

$$\begin{aligned} \alpha(u,v) &= \frac{1}{4}\alpha(u+v,u+v) - \frac{1}{4}\alpha(u-v,u-v) \\ &= \frac{1}{4}\beta(u+v,u+v) - \frac{1}{4}\beta(u-v,u-v) \\ &= \beta(u,v). \end{aligned}$$

For $u := \operatorname{Re}\phi$, $v := \operatorname{Im}\phi$, this gives

$$\begin{aligned} \mathscr{H}\phi(x)^2 &= \alpha(u,u) + 2i\alpha(u,v) - \alpha(v,v) \\ &= \beta(u,u) + 2i\beta(u,v) - \beta(v,v) \\ &= \phi(x)^2 + 2\mathscr{H}(\phi\mathscr{H}\phi)(x). \end{aligned}$$

The assertion follows. □

Cotlar's identity actually holds for any $\phi \in \mathcal{S}(\mathbf{R})$. For a proof, see Proposition 2.23 below.

Using an inductive argument, Cotlar's identity allows us to give an upper bound of $\|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))}$ for p of the form 2^k for $k \in \mathbf{Z}_{\geq 0}$. We have already established that $\|\mathscr{H}\|_{\mathcal{L}(L^2(\mathbf{R}))} = 1 = \cot(\pi/4)$ in Proposition 2.12. For $p \in [2, \infty[$ we write $c_p := \cot(\pi/(2p))$. Now suppose we have established that c_p is an upper bound of $\|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))}$ for $p = 2^k$ for some $k \in \mathbf{Z}_{\geq 0}$. Let $\phi \in C_c^\infty(\mathbf{R})$ satisfy $\|\phi\|_{2p} = 1$. Then, by Cotlar's identity, we have

$$\begin{aligned} \|\mathscr{H}\phi\|_{2p}^2 = \|(\mathscr{H}\phi)^2\|_p &\leq \|\phi^2\|_p + 2\|\mathscr{H}(\phi\mathscr{H}\phi)\|_p \\ &\leq 1 + 2c_p\|\phi\mathscr{H}\phi\|_p \leq 1 + 2c_p\|\mathscr{H}\phi\|_{2p}, \end{aligned}$$

where the last step uses the Cauchy-Schwarz inequality. As the polynomial $t \mapsto t^2 - 2c_p t - 1$ has zeroes $t_\pm = c_p \pm (1+c_p^2)^{1/2}$, we conclude from

$$(\|\mathscr{H}\phi\|_{2p} - t_+)(\|\mathscr{H}\phi\|_{2p} - t_-) = \|\mathscr{H}\phi\|_{2p}^2 - 2c_p\|\mathscr{H}\phi\|_{2p} - 1 \le 0$$

that

$$\|\mathscr{H}\phi\|_{2p} \le t_+ = c_p + (1+c_p^2)^{\frac{1}{2}} = c_{2p},$$

where we have used the trigonometric identity

$$\cot x = \cot 2x + (1 + (\cot 2x)^2)^{\frac{1}{2}}$$

valid for $x \in]0, \pi/4[$. We conclude that $\|\mathscr{H}\|_{\mathcal{L}(L^{2p}(\mathbf{R}))} \le c_{2p}$. Since $2p = 2^{k+1}$, this concludes the inductive step.

As a matter of fact, if we can use other means, such as the Fourier transform, see Corollary 2.21 below, to show that the Hilbert transform extends to an isometry of $L^2(\mathbf{R})$, this yields another proof of Proposition 2.12. Note that this does not give a circular argument, since the proof we gave of Cotlar's Identity only used the fact that $\mathscr{H}$ extends to $L^2(\mathbf{R})$. So far we have solely rested on the theory of complex analysis and we have not used the Fourier transform at all. However, the Fourier transform can be used to help us to sharpen the results we have established so far, which is why we will be working in the distributional setting from now on. See also Appendix B. This emphasizes but one of the many examples of interplay between Fourier analysis and complex analysis, more of which we will encounter in the succeeding sections.

The Hilbert transform is given by convolution with the distribution $\operatorname{PV} 1/t = \operatorname{PV} t/|t|^2$. As it turns out, this is actually a tempered distribution. Temporarily generalizing to $\mathbf{R}^n$, for $j \in \{1, \ldots, n\}$ we define $\operatorname{PV} x_j/|x|^{n+1}$ by

$$\left\langle \operatorname{PV} \frac{x_j}{|x|^{n+1}}, \phi \right\rangle := \lim_{\varepsilon \downarrow 0} \int_{|x| \ge \varepsilon} \frac{x_j}{|x|^{n+1}} \phi(x) \,\mathrm{d}x$$

for $\phi \in \mathcal{S}(\mathbf{R}^n)$. In the following lemma we will show that this is well-defined.

2.18 Lemma. *For all $j \in \{1, \ldots, n\}$ we have $\operatorname{PV} x_j/|x|^{n+1} \in \mathcal{S}'(\mathbf{R}^n)$.*

Proof. We fix $j \in \{1, \ldots, n\}$ and denote by (S^{n-1}, σ) the unit sphere in $\mathbf{R}^n$ with its usual surface measure. Then we observe that

$$\int_{S^{n-1}} \omega_j \,\mathrm{d}\sigma(\omega) = 0 \tag{2.26}$$

by symmetry. Now let $\varepsilon \in]0, 1[$. Then, by employing polar coordinates, it follows that

$$\int_{\varepsilon \le |x| \le 1} \frac{x_j}{|x|^{n+1}} \,\mathrm{d}x = \int_\varepsilon^1 \frac{1}{r} \int_{S^{n-1}} \omega_j \,\mathrm{d}\sigma(\omega) \,\mathrm{d}r = 0.$$

Hence, if $\phi \in \mathcal{S}(\mathbf{R}^n)$, we have

$$\int_{\varepsilon \le |x| \le 1} \frac{x_j}{|x|^{n+1}} \phi(x) \,\mathrm{d}x = \int_{\varepsilon \le |x| \le 1} \frac{x_j}{|x|^{n+1}} (\phi(x) - \phi(0)) \,\mathrm{d}x. \tag{2.27}$$

Recall that the seminorms $(\nu_{j,k})_{j,k\in\mathbf{Z}_{\geq 0}}$ on $\mathcal{S}(\mathbf{R}^n)$ defined by

$$\nu_{j,k}(\psi) = \sup_{x\in\mathbf{R}^n,\, |\alpha|\leq k} (1+|x|)^j |\partial^\alpha \psi(x)|,$$

for $j,k\in\mathbf{Z}_{\geq 0}$, $\psi\in\mathcal{S}(\mathbf{R}^n)$, generate the topology on $\mathcal{S}(\mathbf{R}^n)$. Since

$$|\phi(x)-\phi(0)| \leq \int_0^1 |\partial_t \phi(t\cdot)(x)|\,\mathrm{d}t \leq \int_0^1 \sum_{j=1}^n |\partial_j\phi(tx)x_j|\,\mathrm{d}t \leq n^{\frac{1}{2}}|x|\nu_{0,1}(\phi)$$

for all $x\in\mathbf{R}^n$, we find that

$$\begin{aligned}\int_{|x|\leq 1}\frac{|x_j|}{|x|^{n+1}}|\phi(x)-\phi(0)|\,\mathrm{d}x &\leq n^{\frac{1}{2}}\nu_{0,1}(\phi)\int_{|x|\leq 1}\frac{|x_j|}{|x|^n}\,\mathrm{d}x\\ &= n^{\frac{1}{2}}\nu_{0,1}(\phi)\int_0^1 r^{n-1}r^{-(n-1)}\int_{S^{n-1}}|\omega_j|\,\mathrm{d}\sigma(\omega)\,\mathrm{d}r\\ &= n^{\frac{1}{2}}\nu_{0,1}(\phi)\int_{S^{n-1}}|\omega_j|\,\mathrm{d}\sigma(\omega)<\infty.\end{aligned}$$

Hence, by (2.27) we may conclude that

$$\left|\lim_{\varepsilon\downarrow 0}\int_{\varepsilon\leq|x|\leq 1}\frac{x_j}{|x|^{n+1}}\phi(x)\,\mathrm{d}x\right| = \left|\int_{|x|\leq 1}\frac{x_j}{|x|^{n+1}}(\phi(x)-\phi(0))\,\mathrm{d}x\right| \leq c\nu_{0,1}(\phi) \tag{2.28}$$

for some $c\in\mathbf{R}_+$.

Finally, note that

$$\begin{aligned}\int_{|x|\geq 1}\frac{|x_j|}{|x|^{n+1}}|\phi(x)|\,\mathrm{d}x &\leq \nu_{1,0}(\phi)\int_{|x|\geq 1}\frac{|x_j|}{|x|^{n+2}}\,\mathrm{d}x = \nu_{1,0}(\phi)\int_1^\infty r^{-2}\int_{S^{n-1}}|\omega_j|\,\mathrm{d}\sigma(\omega)\,\mathrm{d}r\\ &= \nu_{1,0}(\phi)\int_{S^{n-1}}|\omega_j|\,\mathrm{d}\sigma(\omega).\end{aligned} \tag{2.29}$$

Thus, by combining (2.28) and (2.29) we have now shown that the limit $\langle \mathrm{PV}\, x_j/|x|^{n+1},\phi\rangle$ exists and that there exist $c,c'\in\mathbf{R}_+$ so that

$$\left|\left\langle \mathrm{PV}\,\frac{x_j}{|x|^{n+1}},\phi\right\rangle\right| \leq c\nu_{0,1}(\phi)+c'\nu_{1,0}(\phi).$$

This proves that $\mathrm{PV}\,x_j/|x|^{n+1}\in\mathcal{S}'(\mathbf{R}^n)$, as desired. □

One can now show that the convolution of $\mathrm{PV}\,x_j/|x|^{n+1}$ with a function $\phi\in\mathcal{S}(\mathbf{R}^n)$ is given by the function

$$\mathrm{PV}\,\frac{x_j}{|x|^{n+1}}*\phi(y) = \lim_{\varepsilon\downarrow 0}\int_{|x-y|\geq\varepsilon}\frac{y_j-x_j}{|y-x|^{n+1}}\phi(x)\,\mathrm{d}x.$$

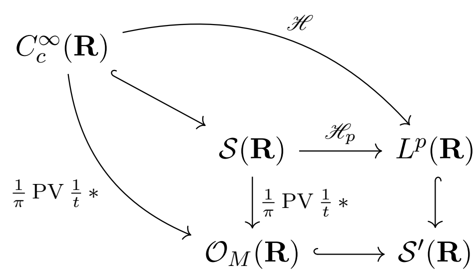
Actually, it follows from Proposition B.41 that these convolution operators, including the Hilbert transform, define continuous linear maps from $\mathcal{S}(\mathbf{R}^n)$ into $\mathcal{O}_M(\mathbf{R}^n)$, the space of moderately increasing functions. Since we are now working with several different extension of the Hilbert transform, it is prudent to check that these notions all coincide. For $p \in]1, \infty[$, denote the extension of $\mathscr{H}$ to $L^p(\mathbf{R})$ by $\mathscr{H}_p$. We should check that for all $\phi \in \mathcal{S}(\mathbf{R})$ we have

$$\mathscr{H}_p \phi(x) = \frac{1}{\pi} \operatorname{PV} \frac{1}{t} * \phi(x) = \lim_{\varepsilon \downarrow 0} \int_{|x-t| \geq \varepsilon} \frac{\phi(t)}{x-t} \, \mathrm{d}t$$

for all $x \in \mathbf{R}$. Fixing $\phi \in \mathcal{S}(\mathbf{R})$, we can use the density of $C_c^\infty(\mathbf{R})$ in $\mathcal{S}(\mathbf{R})$ to find a sequence $(\phi_j)_{j \in \mathbf{N}}$ in $C_c^\infty(\mathbf{R})$ that converges in $\mathcal{S}(\mathbf{R})$ to ϕ. Since $\mathcal{S}(\mathbf{R})$ is continuously included in $L^p(\mathbf{R})$, this sequence also converges to ϕ in $L^p(\mathbf{R})$. Since $\mathscr{H}_p \in \mathcal{L}(L^p(\mathbf{R}))$, the sequence $(\mathscr{H}_p \phi_j)_{j \in \mathbf{N}}$ then converges to $\mathscr{H}_p \phi$ in $L^p(\mathbf{R})$. Since $L^p(\mathbf{R})$ is continuously included in $\mathcal{S}'(\mathbf{R})$, the sequence $(\mathscr{H}_p \phi_j)_{j \in \mathbf{N}}$ also converges in $\mathcal{S}'(\mathbf{R})$ to $\mathscr{H}_p \phi$. On the other hand, since the convolution operator $1/\pi \operatorname{PV} 1/t*$ is a continuous map from $\mathcal{S}(\mathbf{R})$ to $\mathcal{O}_M(\mathbf{R})$, the latter being continuously included in $\mathcal{S}'(\mathbf{R})$, we find that $1/\pi \operatorname{PV} 1/t * \phi_j$ converges to $1/\pi \operatorname{PV} 1/t * \phi$ in $\mathcal{S}'(\mathbf{R})$. Thus, since $\mathcal{S}'(\mathbf{R})$ is Hausdorff, we may take limits in $\mathcal{S}'(\mathbf{R})$ to conclude that

$$\mathscr{H}_p \phi = \lim_{j \to \infty} \mathscr{H}_p \phi_j = \lim_{j \to \infty} \frac{1}{\pi} \operatorname{PV} \frac{1}{t} * \phi_j = \frac{1}{\pi} \operatorname{PV} \frac{1}{t} * \phi,$$

Since $\mathscr{H}_p|_{C_c^\infty(\mathbf{R})} = 1/\pi \operatorname{PV} 1/t * |_{C_c^\infty(\mathbf{R})}$. To clarify, we have shown that the below diagram, where the arrow on the top represents our initial definition of the Hilbert transform, is commutative.



2.19 Definition. Let $U \subseteq \mathbf{R}^n$ be open and $p \in [1, \infty]$. For any $f \in L^\infty(U)$ we define the *multiplication operator* $M_f : L^p(U) \to L^p(U)$ by $M_f g := fg$. We use a similar definition for $L^p_{loc}(U)$. $\diamondsuit$

We note that for any $f \in L^\infty(U)$ we have $M_f^* = M_f$ and $M_f^\dagger = M_{\overline{f}}$.

We normalize our Fourier transform in $\mathbf{R}^n$ so that for $\phi \in \mathcal{S}(\mathbf{R}^n)$ we have

$$\mathscr{F} \phi(\xi) = \int_{\mathbf{R}^n} \phi(x) e^{-2\pi i x \cdot \xi} \, \mathrm{d}x.$$

The following lemma shows that in the Fourier domain, the Hilbert transform is just multiplication by the function $-i \operatorname{sign}$.

2.20 Proposition. *We have*

$$\mathscr{F}(\mathscr{H} f) = M_{-i \operatorname{sign}} \mathscr{F} f$$

for all $f \in L^2(\mathbf{R})$.

Proof. By density of $\mathcal{S}(\mathbf{R})$ in $L^2(\mathbf{R})$ it suffices to consider functions in $\mathcal{S}(\mathbf{R})$. By Proposition B.41 we have

$$\mathscr{F}(\mathscr{H}\phi) = \frac{1}{\pi}\mathscr{F}\left(\mathrm{PV}\,\frac{1}{t} * \phi\right) = \frac{1}{\pi}\mathscr{F}\left(\mathrm{PV}\,\frac{1}{t}\right)\mathscr{F}\phi$$

for all $\phi \in \mathcal{S}(\mathbf{R})$. Thus, it remains to show that

$$\mathscr{F}\left(\mathrm{PV}\,\frac{1}{t}\right) = -\pi i\,\mathrm{sign}\,. \tag{2.30}$$

A direct computation shows that for all $y \in \mathbf{R}_+$ we have $\mathscr{F}(i\,\mathrm{sign}\,e^{-2\pi y|\cdot|}) = Q_y$. Let $\phi \in C_c^\infty(\mathbf{R})$ and define $S\phi \in C_c^\infty(\mathbf{R})$ by $S\phi(t) := \phi(-t)$. Then, by Lemma 2.11, we have

$$\lim_{y\downarrow 0}\langle Q_y, \phi\rangle = \lim_{y\downarrow 0}(Q_y * S\phi)(0) = \mathscr{H}(S\phi)(0) = \left\langle \frac{1}{\pi}\,\mathrm{PV}\,\frac{1}{t}, \phi\right\rangle.$$

Hence, Q_y converges to $1/\pi\,\mathrm{PV}\,1/t$ in $\mathcal{D}'(\mathbf{R})$ as $y \downarrow 0$. Thus, we have

$$\frac{1}{\pi}\,\mathrm{PV}\,\frac{1}{t} = \lim_{y\downarrow 0} Q_y = \lim_{y\downarrow 0}\mathscr{F}(i\,\mathrm{sign}\,e^{-2\pi y|\,\cdot\,|}), \tag{2.31}$$

in $\mathcal{D}'(\mathbf{R})$. Since $i\,\mathrm{sign}\,e^{-2\pi y|\,\cdot\,|}$ converges to $i\,\mathrm{sign}$ in $\mathcal{S}'(\mathbf{R})$ as $y \downarrow 0$, we find, by continuity of the Fourier transform in $\mathcal{S}'(\mathbf{R})$, that

$$\lim_{y\downarrow 0}\mathscr{F}(i\,\mathrm{sign}\,e^{-2\pi y|\,\cdot\,|}) = i\mathscr{F}(\mathrm{sign}), \tag{2.32}$$

in $\mathcal{S}'(\mathbf{R})$. But since $\mathcal{S}'(\mathbf{R})$ is continuously included in $\mathcal{D}'(\mathbf{R})$, the limit in (2.32) also holds in $\mathcal{D}'(\mathbf{R})$. Thus, we may conclude from (2.31), that

$$\frac{1}{\pi}\,\mathrm{PV}\,\frac{1}{t} = i\mathscr{F}(\mathrm{sign}).$$

Fourier inversion then yields (2.30), as desired. □

Assuming one has not yet established L^2-boundedness of the Hilbert transform, one could formulate and prove the result of the proposition as we have for functions in $\mathcal{S}(\mathbf{R})$. Then one obtains:

2.21 Corollary. *The Hilbert transform extends to an isometry of $L^2(\mathbf{R})$.*

Proof. For any $\phi \in \mathcal{S}(\mathbf{R})$ we have

$$\|\mathscr{H}\phi\|_2 = \|\mathscr{F}\mathscr{H}\phi\|_2 = \|M_{-i\,\mathrm{sign}}\mathscr{F}\phi\|_2 = \|\phi\|_2.$$

The assertion follows. □

Another consequence is the following:

2.22 Corollary. *Let $p \in]1,\infty[$. For any $f \in L^p(\mathbf{R})$ we have $\mathscr{H}^2 f = -f$. In other words, $\mathscr{H}$ is a linear automorphism of $L^p(\mathbf{R})$ with inverse $-\mathscr{H}$.*

Proof. Let $\phi \in \mathcal{S}(\mathbf{R})$. Then, since $\mathscr{H}\phi \in L^2(\mathbf{R})$,

$$\mathscr{F}(\mathscr{H}^2\phi) = M^2_{-i\,\mathrm{sign}}\mathscr{F}\phi = \mathscr{F}(-\phi).$$

Thus, $\mathscr{H}^2\phi = -\phi$. By density of $\mathcal{S}(\mathbf{R})$ in $L^p(\mathbf{R})$, this result extends to all functions in $L^p(\mathbf{R})$, as desired. □

We can also give another proof of Cotlar's Identity, extending Proposition 2.17 to functions in $\mathcal{S}(\mathbf{R})$.

2.23 Proposition (Cotlar's Identity)**.** *For any $\phi \in \mathcal{S}(\mathbf{R})$ we have*

$$(\mathscr{H}\phi)^2 = \phi^2 + 2\mathscr{H}(\phi\mathscr{H}\phi).$$

Proof. Let $\phi \in \mathcal{S}(\mathbf{R})$ and $\xi \in \mathbf{R}$, $\xi \neq 0$. Using the fact that for $\psi \in \mathcal{S}(\mathbf{R})$ and $u \in \mathcal{S}'(\mathbf{R})$ we have $\mathscr{F}(\psi u) = \mathscr{F}\psi * \mathscr{F}u$, we note that

$$\begin{aligned}
\mathscr{F}(\phi^2 + 2\mathscr{H}(\phi\mathscr{H}\phi))(\xi) &= (\mathscr{F}\phi * \mathscr{F}\phi)(\xi) - 2i\,\mathrm{sign}(\xi)(\mathscr{F}\phi * M_{-i\,\mathrm{sign}}\mathscr{F}\phi)(\xi) \\
&= (\mathscr{F}\phi * \mathscr{F}\phi)(\xi) + 2\,\mathrm{sign}(\xi)\int_{\mathbf{R}} \mathscr{F}\phi(\xi-\eta)\mathscr{F}\phi(\eta)\,\mathrm{sign}(\eta)\,\mathrm{d}\eta \qquad (2.33) \\
&= (\mathscr{F}\phi * \mathscr{F}\phi)(\xi) + 2\,\mathrm{sign}(\xi)\int_{\mathbf{R}} \mathscr{F}\phi(\eta)\mathscr{F}\phi(\xi-\eta)\,\mathrm{sign}(\xi-\eta)\,\mathrm{d}\eta. \qquad (2.34)
\end{aligned}$$

Thus, by adding (2.33) and (2.34) and dividing by 2, we obtain

$$\mathscr{F}(\phi^2 + 2\mathscr{H}(\phi\mathscr{H}\phi))(\xi) = \int_{\mathbf{R}} \mathscr{F}\phi(\eta)\mathscr{F}\phi(\xi-\eta)(1 - \mathrm{sign}(\xi)(\mathrm{sign}(\eta) - \mathrm{sign}(\xi-\eta)))\,\mathrm{d}\eta. \qquad (2.35)$$

A proof by cases reveals that

$$1 - \mathrm{sign}(\xi)(\mathrm{sign}(\eta) - \mathrm{sign}(\xi-\eta)) = -\,\mathrm{sign}(\eta)\,\mathrm{sign}(\xi-\eta) = (-i\,\mathrm{sign}(\eta))(-i\,\mathrm{sign}(\xi-\eta)),$$

for all $\eta \in \mathbf{R}$ with $\eta \neq 0$ and $\eta \neq \xi$. Hence, by (2.35), we have

$$\mathscr{F}(\phi^2 + 2\mathscr{H}(\phi\mathscr{H}\phi))(\xi) = \int_{\mathbf{R}} M_{-i\,\mathrm{sign}}\mathscr{F}\phi(\eta)M_{-i\,\mathrm{sign}}\mathscr{F}\phi(\xi-\eta)\,\mathrm{d}\eta = \mathscr{F}((\mathscr{H}\phi)^2)(\xi),$$

where the last equality follows from the formula $\mathscr{F}(f^2) = \mathscr{F}f * \mathscr{F}f$ valid for $f \in L^\infty(\mathbf{R})$ satisfying $\mathscr{F}f \in L^1(\mathbf{R})$ where $f = \mathscr{H}\phi$. To see this, we note that indeed $\hat{f} := \mathscr{F}f = M_{-i\,\mathrm{sign}}\mathscr{F}\phi \in L^1(\mathbf{R})$. Then $f = \mathscr{F}^{-1}\hat{f} \in L^\infty(\mathbf{R})$ and $\mathscr{F}^{-1}(\hat{f} * \hat{f}) = (\mathscr{F}^{-1}\hat{f})^2 = f^2$ by Theorem B.32. The assertion follows. □

Next, we will extend the result of Lemma 2.11.

2.24 Lemma. *Let $p \in]1, \infty[$ and $f \in L^p(\mathbf{R})$. Then*

$$\lim_{y \downarrow 0} Q_y * f = \mathscr{H}f,$$

where the limit is in $L^p(\mathbf{R})$.

Proof. Let $p \in]1,\infty[$. Fix $y \in \mathbf{R}_+$ and $x \in \mathbf{R}$. For all $t \in \mathbf{R}$ and $f \in L^p(\mathbf{R})$ we note that

$$\frac{|x-t|}{|x-t|^2+y^2} \le \frac{1}{y^2} \quad \text{if } |x-t| \le 1 \text{ and } \quad \frac{|x-t|}{|x-t|^2+y^2} \le \frac{1}{|x-t|} \quad \text{if } |x-t| \ge 1.$$

Hence,

$$\begin{aligned}\int_{\mathbf{R}} |f(t)| \frac{|x-t|}{|x-t|^2+y^2}\,\mathrm{d}t &\le \int_{|x-t|\le 1} \frac{|f(t)|}{y^2}\,\mathrm{d}t + \int_{|x-t|\ge 1} \frac{|f(t)|}{|x-t|}\,\mathrm{d}t \\ &\le \frac{2^{\frac{1}{p'}}\|f\|_p}{y^2} + \|f\|_p \int_{|t|\ge 1} |t|^{-p'}\,\mathrm{d}t \\ &= \left(\frac{2^{\frac{1}{p'}}}{y^2} + \frac{2}{p'-1}\right)\|f\|_p,\end{aligned}$$

by Hölders inequality. This means that $Q_y * f$ is well-defined for all $f \in L^p(\mathbf{R})$ and that the map $f \mapsto Q_y * f(x)$ is continuous from $L^p(\mathbf{R})$ to $\mathbf{C}$.

The remainder of this proof is based on the identity

$$Q_y * f = P_y * \mathscr{H} f, \tag{2.36}$$

valid for any $f \in L^p(\mathbf{R})$. Indeed, if this identity holds then the result follows from Lemma 2.11.

Assume for the moment we have shown that (2.36) holds for all $f \in \mathcal{S}(\mathbf{R})$. If $f \in L^p(\mathbf{R})$ is arbitrary, then we can pick a sequence $(\phi_j)_{j\in\mathbf{N}}$ in $\mathcal{S}(\mathbf{R})$ that converges to f in $L^p(\mathbf{R}^n)$. Hence, $Q_y * \phi_j \to Q_y * f$ pointwise as $j \to \infty$. By Minkowski's inequality for convolutions, see Corollary A.5, the right-hand side of (2.36) as a function of f is continuous as a map from $L^p(\mathbf{R})$ to $L^p(\mathbf{R})$. Hence, $P_y * \mathscr{H}\phi_j \to P_y * \mathscr{H} f$ as $j \to \infty$. But this implies there is some a.e. convergent subsequence of $(P_y * \mathscr{H}\phi_j)_{j\in\mathbf{N}}$ with limit $P_y * \mathscr{H} f$. Thus, taking a.e. limits, we conclude from the fact that $Q_y * \phi_j = P_y * \mathscr{H}\phi_j$ for all $j \in \mathbf{N}$ that (2.36) is valid for all $f \in L^p(\mathbf{R})$.

It remains to prove that (2.36) is valid for all $f \in \mathcal{S}(\mathbf{R})$. A direct computations show that $\mathscr{F}(e^{-2\pi y|\,\cdot\,|}) = P_y$ and $\mathscr{F}(i \operatorname{sign} e^{-2\pi y|\,\cdot\,|}) = Q_y$. Fourier inversion then yields

$$\mathscr{F} P_y = e^{-2\pi y|\,\cdot\,|}, \quad \mathscr{F} Q_y = -i \operatorname{sign} e^{-2\pi y|\,\cdot\,|}.$$

Let $f \in \mathcal{S}(\mathbf{R})$. Then

$$\mathscr{F}(Q_y * f) = \mathscr{F} Q_y \mathscr{F} f = -i \operatorname{sign} e^{-2\pi y|\,\cdot\,|} \mathscr{F} f = \mathscr{F} P_y \mathscr{F}(\mathscr{H} f) = \mathscr{F}(P_y * \mathscr{H} f)$$

so that the assertion follows by Fourier inversion, where the first equality follows from the formula $\mathscr{F}(u * \phi) = \mathscr{F} u \mathscr{F} \phi$ valid for $u \in \mathcal{S}'(\mathbf{R})$ and $\phi \in \mathcal{S}(\mathbf{R})$ applied to $u = Q_y$ and $\phi = f$, while the last equality follows from the formula $\mathscr{F}(g * h) = \mathscr{F} g \mathscr{F} h$ valid for $g \in L^1(\mathbf{R})$ and $h \in L^2(\mathbf{R})$ applied to $g = P_y$ and $h = \mathscr{H} f$, see Proposition B.41 and Theorem B.32. This proves the result. □

Having shown this, we can now prove Theorem 2.4

Proof of Theorem 2.4. By Lemma 2.16 we have $2f = Cg$. Hence,

$$\begin{aligned} 2\operatorname{Im} f(x+iy) &= \operatorname{Im}((P_y * g)(x) + i(Q_y * g)(x)) \\ &= (P_y * \operatorname{Im} g)(x) + (Q_y * \operatorname{Re} g)(x) \end{aligned} \tag{2.37}$$

for all $z = x + iy \in \mathbf{H}$. Since

$$\lim_{y \downarrow 0} (P_y * \operatorname{Im} g) + (Q_y * \operatorname{Re} g) = \operatorname{Im} g + \mathscr{H}(\operatorname{Re} g)$$

in $L^p(\mathbf{R})$ by Lemma 2.11 and Lemma 2.24, the sequence $((P_{1/j} * \operatorname{Im} g) + (Q_{1/j} * \operatorname{Re} g))_{j \in \mathbf{N}}$ has an a.e. convergent subsequence with limit $\operatorname{Im} g + \mathscr{H}(\operatorname{Re} g)$. By combining this result with (2.5), we conclude from (2.37) that

$$2\operatorname{Im} g = \operatorname{Im} g + \mathscr{H}(\operatorname{Re} g).$$

The assertion follows. □

With this result, we will work towards the proof of operator norm equalities in Theorem 2.3. We will work with the outline provided by [Gr, Exercise 4.1.13] using results from the article [Pi] by Pichorides, which gave the original proof of this result in 1972. On account of Proposition 2.6 it suffices to check the cases where $p \in]1, 2]$. Since we have already established the case where $p = 2$, we may assume that $p \in]1, 2[$. First we will establish the lower bound

$$\|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))} \geq \tan \frac{\pi}{2p}. \tag{2.38}$$

We define $\phi : \mathbf{C} \backslash \{1\} \to \mathbf{C}$ to be the conformal mapping

$$\phi(z) := i\frac{z+1}{z-1} = \frac{2y}{|z-1|^2} + i\frac{|z|^2 - 1}{|z-1|^2}.$$

Now let $U \subseteq \mathbf{C}$ be the open subset of $\mathbf{C}$ obtained by deleting the non-positive real numbers from $\mathbf{C}$. The function ϕ maps $\mathbf{H}$ into the open right half plane of $\mathbf{C}$. In particular, this means that $\phi(\mathbf{H}) \subseteq U$. Any $z \in U$ can be written as $z = re^{it}$ with $r \in \mathbf{R}_+$ and $t \in]-\pi, \pi[$. We let $\log : U \to \mathbf{C}$ be the holomorphic function satisfying $\log z = \log r + it$, where $\log r$ denotes the natural logarithm of r. See Example C.24. For

$$\gamma \in \left] \frac{\pi}{2p'}, \frac{\pi}{2p} \right[, \tag{2.39}$$

this allows us to define an analytic function

$$f : \mathbf{H} \to \mathbf{C}, \quad f(z) := \frac{\phi(z)^{\frac{2\gamma}{\pi}}}{z+1} := \frac{e^{\frac{2\gamma}{\pi} \log \phi(z)}}{z+1}.$$

We wish to apply Theorem 2.4 to this function. Since $|\phi(z)| \to 1$ as $|z| \to \infty$, we find that $|zf(z)| \to 1$ as $|z| \to \infty$. This implies that f satisfies (2.6).

Next, note that for $x \in \mathbf{R} \backslash \{\pm 1\}$ we can write $\phi(x) = r(x)e^{it(x)}$ where $r(x) = |x+1|/|x-1|$ and

$$t(x) = \begin{cases} \frac{\pi}{2} & \text{if } |x| > 1 \\ -\frac{\pi}{2} & \text{if } |x| < 1. \end{cases}$$

This implies that if we set $g(x) := \lim_{y\downarrow 0} f(x+iy)$ for $x \in \mathbf{R}\backslash\{\pm 1\}$, then

$$g(x) = \frac{1}{x+1}\left|\frac{x+1}{x-1}\right|^{\frac{2\gamma}{\pi}} e^{i\frac{2\gamma}{\pi}t(x)}.$$

Note that

$$|g(x)|^p = |x+1|^{p\left(\frac{2\gamma}{\pi}-1\right)}|x-1|^{-p\frac{2\gamma}{\pi}}.$$

For large $|x|$, the function $|g(x)|^p$ behaves like $|x+1|^{-p}$ which is integrable in a neighborhood of infinity, since $-p < -1$. Moreover, $|g(x)|^p$ behaves like $|x+1|^{p(2\gamma/\pi-1)}$ for x near -1 and like $|x-1|^{-2p\gamma/\pi}$ for x near 1. Since our assumptions on γ imply that $-1 < p(2\gamma/\pi - 1)$ and $-1 < -2p\gamma/\pi$, we may conclude that $|g(x)|^p$ is integrable for x near ± 1. By combining these results we have found that $g \in L^p(\mathbf{R})$. Thus, we may conclude from Theorem 2.4 that $\mathscr{H}(\operatorname{Re} g) = \operatorname{Im} g$. Noting that

$$\operatorname{Re} g(x) = \frac{1}{x+1}\left|\frac{x+1}{x-1}\right|^{\frac{2\gamma}{\pi}} \cos\gamma$$

and

$$\operatorname{Im} g(x) = \begin{cases} \frac{1}{x+1}\left|\frac{x+1}{x-1}\right|^{\frac{2\gamma}{\pi}} \sin\gamma & \text{if } |x| > 1 \\ -\frac{1}{x+1}\left|\frac{x+1}{x-1}\right|^{\frac{2\gamma}{\pi}} \sin\gamma & \text{if } |x| < 1, \end{cases}$$

this means that

$$\|\mathscr{H}(\operatorname{Re} g)\|_p = \|\operatorname{Im} g\|_p = \tan\gamma \|\operatorname{Re} g\|_p.$$

Hence,

$$\|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))} \geq \tan\gamma$$

for all γ satisfying (2.39). From this we may conclude that the lower bound (2.38) holds.

In view of our discussion after Proposition 2.17, we have now shown the operator norm equality in Theorem 2.3 for $p \in]1,\infty[$ of the form $p = 2^k$ and $p = 2^k/(2^k-1)$ for $k \in \mathbf{Z}_{\geq 0}$.

The remainder of this subsection will be dedicated to proving the upper bound

$$\|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))} \leq \tan\frac{\pi}{2p} \tag{2.40}$$

for $p \in]1,2[$. We will actually show that for such p we have

$$\|\mathscr{H}f\|_p \leq \tan\frac{\pi}{2p}\|f\|_p \quad \text{for all real-valued } f \in L^p(\mathbf{R}). \tag{2.41}$$

We sketch a way to conclude the general complex-valued result (2.40) from (2.41).

2.25 Lemma. *Let $p \in]1,\infty[$ and $\mathscr{L} \in \mathcal{L}(L^p(\mathbf{R}))$. Suppose $\mathscr{L}$ maps real-valued functions to real valued functions and suppose there is some $c \in \mathbf{R}_+$ so that*

$$\|\mathscr{L}f\|_p \leq c\|f\|_p \quad \textit{for all real-valued } f \in L^p(\mathbf{R}).$$

Then

$$\|\mathscr{L}f\|_p \leq c\|f\|_p$$

for all $f \in L^p(\mathbf{R})$.

Proof. Define $\gamma : \mathbf{C} \to \mathbf{R}$ by $\gamma(z) := e^{-\pi|z|^2}$. Then γ defines a measure $\mathrm{d}\gamma$ on $\mathbf{C}$ through $\mathrm{d}\gamma(z) = \gamma(z)\,\mathrm{d}z$. We denote by $\mathrm{d}\gamma_{\mathbf{R}}$ the measure on $\mathbf{R}$ obtained as the image measure of $\mathrm{d}\gamma$ under the projection $z \mapsto \operatorname{Re} z$. Since $\int_{\mathbf{R}} e^{-\pi y^2}\,\mathrm{d}y = 1$, one finds that $\mathrm{d}\gamma_{\mathbf{R}}(x) = e^{-\pi x^2}\,\mathrm{d}x$. If $w \in \mathbf{C}$ satisfies $|w| = 1$, then, since γ is rotationally invariant, we find that the image measure of $\mathrm{d}\gamma$ under $z \mapsto \operatorname{Re}(\overline{w}z)$ is also $\mathrm{d}\gamma_{\mathbf{R}}$. This implies that

$$k := \pi^{-\frac{p+1}{2}}\Gamma\left(\frac{p+1}{2}\right) = \int_{\mathbf{R}} |t|^p\,\mathrm{d}\gamma_{\mathbf{R}}(t) = \int_{\mathbf{C}} |\operatorname{Re}(\overline{w}z)|^p\,\mathrm{d}\gamma(z).$$

Thus, for any $w \in \mathbf{C}$ we may conclude that

$$\int_{\mathbf{C}} |\operatorname{Re}(\overline{w}z)|^p\,\mathrm{d}\gamma(z) = k|w|^p.$$

Hence, for any $g \in L^p(\mathbf{R})$ we have

$$k\|g\|_p^p = \int_{\mathbf{C}}\int_{\mathbf{R}} |\operatorname{Re}(\overline{g(t)}z)|^p\,\mathrm{d}t\,\mathrm{d}\gamma(z). \tag{2.42}$$

Let $f \in L^p(\mathbf{R})$. Since $\mathscr{L}$ maps real-valued functions to real-valued functions we have

$$\operatorname{Re}(\overline{\mathscr{L}f(t)}z) = \mathscr{L}(\operatorname{Re}(\overline{f}z))(t)$$

for all $t \in \mathbf{R}$ and $z \in \mathbf{C}$. Hence, by (2.42),

$$\begin{aligned} k\|\mathscr{L}f\|_p^p &= \int_{\mathbf{C}} \|\mathscr{L}(\operatorname{Re}\overline{f}z)\|_p^p\,\mathrm{d}\gamma(z) \le c^p \int_{\mathbf{C}} \|(\operatorname{Re}\overline{f}z)\|_p^p\,\mathrm{d}\gamma(z) \\ &= c^p \int_{\mathbf{C}}\int_{\mathbf{R}} |\operatorname{Re}(\overline{f(t)}z)|^p\,\mathrm{d}t\,\mathrm{d}\gamma(z) = c^p k\|f\|_p^p. \end{aligned}$$

The assertion follows. □

The proof we give of (2.41) relies on a certain subharmonic function.

2.26 Definition. Let $p \in]1, 2]$. We define $f_p : \mathbf{C} \to \mathbf{C}$ by $f_p(0) := 0$ and

$$f_p(z) := \operatorname{Re}((|x| + iy)^p) = \operatorname{Re} e^{p\log(|x|+iy)} = |z|^p \cos(p\arg(|x| + iy))$$

for $z = x + iy \neq 0$, where the argument is taken between $-\pi/2$ and $\pi/2$. ♢

The importance of this function becomes clear in the following remarkable inequality:

2.27 Proposition. *Let $p \in]1, 2]$. Then*

$$\frac{\left(\sin\frac{\pi}{2p}\right)^{p-1}}{\cos\frac{\pi}{2p}} f_p(a + bi) \le \left(\tan\frac{\pi}{2p}\right)^p |a|^p - |b|^p \tag{2.43}$$

for all $a, b \in \mathbf{R}$.

Our proof is based on [Pi, Lemma 2.1]. Note that this inequality is actually an equality when $p = 2$.

Proof. We set

$$A_p := \left(\tan \frac{\pi}{2p}\right)^p, \quad B_p := \frac{\left(\sin \frac{\pi}{2p}\right)^{p-1}}{\cos \frac{\pi}{2p}}.$$

The proof uses the inequality

$$|\sin x|^p \leq A_p(\cos x)^p - B_p \cos px, \tag{2.44}$$

valid for $x \in [-\pi/2, \pi/2]$.

Assume for the moment that we have shown validity of (2.44). The inequality (2.43) is immediate for $a = b = 0$. Now let $a, b \in \mathbf{R}$ not both be 0 and $x := \arg(|a| + bi)$. We then find that

$$\sin x = \frac{b}{|a + bi|}, \quad \cos x = \frac{|a|}{|a + bi|}.$$

Hence,

$$\frac{|b|^p}{|a + bi|^p} \leq A_p \frac{|a|^p}{|a + bi|^p} - B_p \cos(p \arg(|a| + bi))$$

by (2.44). Rearranging the terms yields (2.43), as desired.

To prove (2.44), we note that there is equality when $p = 2$. Hence we may assume $p \in]1, 2[$. Moreover, since both sides are even functions in x, we need only consider the cases where $x \in [0, \pi/2]$. But then, by continuity, it suffices to consider the cases where $x \in]0, \pi/2[$. We define

$$f : \left]0, \frac{\pi}{2}\right[\to \mathbf{R}, \quad f(x) := \frac{(\sin x)^p + B_p \cos px}{(\cos x)^p}.$$

If we then set

$$g : \left]0, \frac{\pi}{2}\right[\to \mathbf{R}, \quad g(x) := 1 - B_p \frac{\sin((p-1)x)}{(\sin x)^{p-1}},$$

then

$$f'(x) = p \frac{(\sin x)^{p-1}}{(\cos x)^{p+1}} g(x).$$

We note that, since $2 - p \in]0, 1[$,

$$\begin{aligned} g'(x) &= -B_p(p-1)\cos((p-1)x)(\sin x)^{1-p} + B_p(p-1)\sin((p-1)x)\cos x(\sin x)^{-p} \\ &= -B_p(p-1)\frac{\sin((2-p)x)}{(\sin x)^p} < 0, \end{aligned}$$

meaning that g is strictly decreasing. This allows us to conclude that g and hence f' has a unique zero at $\pi/(2p)$. An application of de L'Hôpital's rule shows, using $p \in]1, 2[$, that $\lim_{x \downarrow 0} g(x) = 1$. This means that for small enough $x \in]0, \pi/2[$ we have $f'(x) > 0$. Moreover, since $f(x) \to -\infty$ as $x \uparrow \pi/2$ we may conclude that f attains a global maximum at $\pi/(2p)$. Hence,

$$\frac{(\sin x)^p + B_p \cos px}{(\cos x)^p} \leq f\left(\frac{\pi}{2p}\right) = A_p,$$

for all $x \in]0, \pi/2[$. This proves (2.44). The assertion follows. $\square$

In view of this proposition, it now suffices to show that

$$\int_{\mathbf{R}} f_p(|\phi(x)| + i\mathscr{H}\phi(x))\,\mathrm{d}x \geq 0$$

for all $\phi \in C_c^\infty(\mathbf{R})$ to conclude (2.40). This will follow from the fact that f_p is a subharmonic function.

2.28 Definition. Let $U \subseteq \mathbf{C}$ be open and let $u : U \to \mathbf{R}$ be a continuous function. We call u *subharmonic* if it satisfies the *mean-value property*

$$u(z_0) \leq \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + re^{it})\,\mathrm{d}t$$

for all $z_0 \in U$, $r \in \mathbf{R}_+$ so that $\overline{D}(z_0; r) \subseteq U$. ♢

We will use the following characterization:

2.29 Lemma. *Let $U \subseteq \mathbf{C}$ be open and let $u : U \to \mathbf{R}$ be a continuous function. The following are equivalent:*

(i) u is subharmonic in U;

(ii) for every $z_0 \in U$ there is an $r_0 \in \mathbf{R}_+$ so that $\overline{D}(z_0; r_0) \subseteq U$ and whenever $0 < r < r_0$, we have

$$u(z_0) \leq \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + re^{it})\,\mathrm{d}t.$$

A discussion on subharmonic functions can be found in Appendix C and, in particular, Lemma 2.29 is a consequence of Theorem C.38.

2.30 Lemma. *Let $p \in]1, 2]$. Then f_p is subharmonic in $\mathbf{C}$.*

Our proof follows the proof of [Pi, Lemma 3.5].

Proof. In the half planes $\{x + iy \in \mathbf{C} \mid x \in \mathbf{R}_+\}$ and $\{x + iy \in \mathbf{C} \mid -x \in \mathbf{R}_+\}$ the function f_p coincides with the real part of the holomorphic functions z^p and $(-z)^p$ respectively and is thus harmonic there. Hence, f_p certainly satisfies the mean-value property at any point in these half planes. It remains to check the purely imaginary points in $\mathbf{C}$.

We first check that f_p satisfies the mean-value property at discs around origin. Let $r \in \mathbf{R}_+$. Then, noting that

$$f_p(re^{it}) = \begin{cases} r^p \cos pt & \text{if } t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ r^p \cos p(\pi - t) & \text{if } t \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right], \end{cases} \tag{2.45}$$

we find that, since $p \in]1, 2]$,

$$\frac{1}{2\pi}\int_0^{2\pi} f_p(re^{it})\,\mathrm{d}t = \frac{2r^p}{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos pt\,\mathrm{d}t = \frac{2r^p \sin p\frac{\pi}{2}}{\pi p} \geq 0 = f_p(0),$$

as desired.

We denote by $U \subseteq \mathbf{C}$ the set of complex numbers minus the non-positive real numbers, and define $u : U \to \mathbf{R}$ as the real part of the holomorphic function z^p. Then u is harmonic in U, see Proposition C.34, and coincides with f_p on the closed right half plane. Let $z = re^{it}$ with $r \in \mathbf{R}_+$ and $t \in [\pi/2, \pi]$. Then, by (2.45), we have

$$f_p(z) - u(z) = r^p \cos p(\pi - t) - r^p \cos pt = 2r^p \sin p\left(t - \frac{\pi}{2}\right) \sin p\frac{\pi}{2} \geq 0.$$

If $z = re^{it}$ with $r \in \mathbf{R}_+$ and $t \in [\pi, 3\pi/2]$, then

$$f_p(z) - u(z) = r^p \cos p(\pi + t) - r^p \cos pt = -2r^p \sin p\left(t + \frac{\pi}{2}\right) \sin p\frac{\pi}{2} \geq 0.$$

Hence, we have found that $u(z) \leq f_p(z)$ for all $z \in U$. This means that if $y \in \mathbf{R}\backslash\{0\}$ and $0 < r < |y|$, then, since u is harmonic in U,

$$f_p(iy) = u(iy) = \frac{1}{2\pi}\int_0^{2\pi} u(iy + re^{it})\,\mathrm{d}t \leq \frac{1}{2\pi}\int_0^{2\pi} f_p(iy + re^{it})\,\mathrm{d}t.$$

The assertion follows. □

2.31 Corollary. *Let $u \in C_c^\infty(\mathbf{R})$ be real-valued and let $p \in]1, 2]$. Then $f_p \circ Cu : \mathbf{H} \to \mathbf{C}$ is subharmonic.*

Proof. Since $Cu : \mathbf{H} \to \mathbf{C}$ is holomorphic, this follows immediately from Lemma C.40. □

As anounced, this result will be used to prove the following:

2.32 Lemma. *Let $p \in]1, 2]$. Then we have*

$$\int_{\mathbf{R}} f_p(u(x) + i\mathscr{H}u(x))\,\mathrm{d}x \geq 0, \tag{2.46}$$

for all real-valued $u \in C_c^\infty(\mathbf{R})$.

Proof. First we will check that $f_p(u + i\mathscr{H}u)$ is integrable over $\mathbf{R}$. Since $u \in C_c^\infty(\mathbf{R}) \subseteq L^p(\mathbf{R})$, we also have $\mathscr{H}u \in L^p(\mathbf{R})$. Noting that $|f_p(u(x) + i\mathscr{H}u(x))| \leq |u(x) + i\mathscr{H}u(x)|^p$ for all $x \in \mathbf{R}$, we may indeed conclude that $f_p(u + i\mathscr{H}u) \in L^1(\mathbf{R})$.

Next, let $r \in \mathbf{R}_+$ and let (C_r, σ_r) denote the circle in $\mathbf{C}$ of radius r, centered at ir, equipped with its usual surface measure σ_r. Let $0 < r' < r$. Then, by Corollary 2.31, we have

$$f_p(Cu(ir)) \leq \frac{1}{2\pi}\int_0^{2\pi} f_p(Cu(ir + r'e^{it}))\,\mathrm{d}t. \tag{2.47}$$

As the convergence in 2.11 is uniform in x, we can conclude that Cu extends continuously to the closure $\overline{\mathbf{H}}$ of $\mathbf{H}$ in $\mathbf{C}$ by declaring that it is equal to $u + i\mathscr{H}u$ on the real line. In particular this means that Cu is bounded on any compact set in $\overline{\mathbf{H}}$. This justifies letting $r' \uparrow r$ in (2.47) to conclude that

$$2\pi r f_p(Cu(ir)) \leq \int_0^{2\pi} r f_p(Cu(ir + re^{it}))\,\mathrm{d}t = \int_{C_r} f_p(Cu(\omega))\,\mathrm{d}\sigma(\omega). \tag{2.48}$$

The idea is now to let $r \to \infty$ in (2.48) to conclude (2.46).

By Lemma 2.14 we can find $R, c \in \mathbf{R}_+$ so that whenever $|z| \geq R$, we have $|Cu(z)| \leq c/|z|$. Then, since $p > 1$, we have

$$|rf_p(Cu(ir))| \leq r|Cu(ir)|^p \leq \frac{rc^p}{|ir|^p} = \frac{c^p}{r^{p-1}} \to 0 \quad \text{as } R \leq r \to \infty,$$

showing that the left-hand side of (2.48) tends to 0 as $r \to \infty$.

For the right-hand side of (2.48), we will parameterize the circle C_r by projecting it to the real line through its north pole $2ir$. For any $t \in \mathbf{R}$, the unique intersection point of the line through $2ir$ and t with $C_r \backslash \{2ir\}$ is given by

$$\gamma_r(t) = \frac{4r^2 t}{4r^2 + t^2} + i\frac{2rt^2}{4r^2 + t^2},$$

from which we compute

$$\gamma_r'(t) = 4r^2 \frac{4r^2 - t^2}{(4r^2 + t^2)^2} + 4r^2 i \frac{4rt}{(4r^2 + t^2)^2}.$$

This means that for all $t \in \mathbf{R}$ we have

$$|\gamma_r(t)| = |t| \left(\frac{4r^2}{4r^2 + t^2}\right)^{\frac{1}{2}}, \quad |\gamma_r'(t)| = \frac{4r^2}{4r^2 + t^2}. \tag{2.49}$$

and

$$\lim_{r\to\infty} \gamma_r(t) = t, \quad \lim_{r\to\infty} |\gamma_r'(t)| = 1, \quad \text{for all } t \in \mathbf{R}. \tag{2.50}$$

The former limit implies that if we take $|t| \geq R + 1$, then for some large enough $r \in \mathbf{R}_+$ we have $|\gamma_r(t)| \geq R$. Hence, there is some $M \in \mathbf{R}_+$ so that whenever $r \geq M$ and $|t| \geq M$, we have, by (2.49), that

$$|\gamma_r'(t)||f_p(Cu(\gamma_r(t)))| \leq c^p \frac{|\gamma_r'(t)|}{|\gamma_r(t)|^p} = \frac{c^p}{|t|^p} \left(\frac{4r^2}{4r^2 + t^2}\right)^{1-\frac{p}{2}} \leq \frac{c^p}{|t|^p},$$

where we used the fact that $p \leq 2$ implies that $1 - p/2 \geq 0$. Since $p > 1$, the function $1/|t|^p$ is integrable over the set where $|t| \geq M$. This justifies the use of Lebesgue's Dominated Convergence Theorem to conclude from (2.50) that

$$\lim_{r\to\infty} \int_{C_r} f_p(Cu(\omega))\,\mathrm{d}\sigma(\omega) = \lim_{r\to\infty} \int_{\mathbf{R}} |\gamma_r'(t)| f_p(Cu(\gamma_r(t)))\,\mathrm{d}t = \int_{\mathbf{R}} f_p(u(x) + i\mathscr{H}u(x))\,\mathrm{d}x.$$

The result now follows by letting $r \to \infty$ in (2.48). □

Now, by combining Proposition 2.27 and Lemma 2.32, we conclude that for all $p \in]1, 2]$ and all real-valued $u \in C_c^\infty(R)$ we have

$$0 \leq \frac{\left(\sin\frac{\pi}{2p}\right)^{p-1}}{\cos\frac{\pi}{2p}} \int_{\mathbf{R}} f_p(u(t) + i\mathscr{H}u(t))\,\mathrm{d}t \leq \left(\tan\frac{\pi}{2p}\right)^p \|u\|_p - \|\mathscr{H}u\|_p.$$

The inequality (2.41) now follows from the density in $L^p(\mathbf{R})$ of the space of real-valued functions in $\in C_c^\infty(\mathbf{R})$ in the space of real-valued functions in $L^p(\mathbf{R})$. This concludes the proof of Theorem 2.3.

2.3 The Riesz Transforms and the Beurling-Ahlfors Transform

The method of rotations uses the Hilbert transform and its extensions in $\mathcal{L}(L^p(\mathbf{R}))$ for $p \in]1,\infty[$ to show that certain singular integral operators in $\mathbf{R}^n$ also have extensions in $\mathcal{L}(L^p(\mathbf{R}^n))$ for $p \in]1,\infty[$. Since we will mainly be working in two dimensions when working with the Beurling-Ahlfors transform, this will be our setting now as well.

While we denote our coordinates in the planar domain by $z = x + iy$, we will denote our coordinates in the Fourier domain by $\zeta = \xi + i\eta$. In an attempt to generalize the Hilbert transform to operators on $\mathbf{C}$ with similar properties, we wish to define the following:

2.33 Definition. The first and second *Riesz transforms* $\mathscr{R}_1$ and $\mathscr{R}_2$ are defined as the convolution operators

$$\mathscr{R}_1\phi := \frac{1}{2\pi}\,\mathrm{PV}\,\frac{x}{|z|^3} * \phi, \frac{1}{2\pi} \quad \mathscr{R}_2\phi := \mathrm{PV}\,\frac{y}{|z|^3} * \phi,$$

where $\phi \in \mathcal{S}(\mathbf{R})$. ◊

These operators are well-defined by Lemma 2.18. Just like with the Hilbert transform, the Riesz transforms continuously map $\mathcal{S}(\mathbf{C})$ to $\mathcal{O}_M(\mathbf{C})$, where

$$\mathscr{R}_1\phi(w) = \frac{1}{2\pi}\lim_{\varepsilon\downarrow 0}\int_{|w-z|\geq\varepsilon}\frac{\mathrm{Re}(w-z)}{|w-z|^3}\phi(z)\,\mathrm{d}z, \quad \mathscr{R}_2\phi(w) = \frac{1}{2\pi}\lim_{\varepsilon\downarrow 0}\int_{|w-z|\geq\varepsilon}\frac{\mathrm{Im}(w-z)}{|w-z|^3}\phi(z)\,\mathrm{d}z.$$

Note that for the Hilbert transform we used the normalization $1/\pi$. Usually the Riesz transforms in $\mathbf{R}^n$ are defined as convolution with the distributions $\Gamma((n+1)/2)/\pi^{(n+1)/2}\,\mathrm{PV}\,x_j/|x|^{n+1}$ for $j \in \{1,\dots,n\}$, which gives our normalization with $1/(2\pi)$ here.

We now state the main objectives of this subsection.

2.34 Theorem. *Let $p \in]1,\infty[$. The Riesz transforms have extensions in $\mathcal{L}(L^p(\mathbf{C}))$ satisfying the estimate $\|\mathscr{R}_j\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq \|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))}$ for $j \in \{1,2\}$.*

Showing this will involve what is called the method of rotations.

Additionally, we will wish to compute the Fourier transform of the Riesz transforms.

2.35 Theorem. *We have*

$$\mathscr{F}(\mathscr{R}_1 f) = M_{-i\frac{\xi}{|\zeta|}}\mathscr{F}f, \quad \mathscr{F}(\mathscr{R}_2 f) = M_{-i\frac{\eta}{|\zeta|}}\mathscr{F}f,$$

for all $f \in L^2(\mathbf{C})$.

The main reason we wish to establish these results, is because we will define the Beurling-Ahlfors transform in terms of the Riesz transforms to establish its L^p boundedness through Theorem 2.34. Then, to see that this definition coincides with usual integral representation of the Beurling-Ahlfors transform, we need only show that their Fourier transforms coincide. After we have established our main results, we will use the remainder of this subsection to discuss properties of the Beurling-Ahlfors transform.

First we will concern ourselves with the proof of Theorem 2.34. For $\alpha \in \mathbf{R}$ we denote by $r_\alpha : \mathbf{C} \to \mathbf{C}$ the rotation $z \mapsto e^{i\alpha}z$. Then r_α defines a map $r_\alpha^* : L^0(\mathbf{C}) \to L^0(\mathbf{C})$, by $r_\alpha^* f = f \circ r_\alpha$. In particular, r_α^* restricts to an isometric linear isomorphism of $L^p(\mathbf{C})$ for all $p \in [1,\infty]$ with inverse $r_{-\alpha}^*$.

2.36 Definition. Let $\alpha \in \mathbf{R}$. We define the *angular Hilbert transform* $\mathscr{H}_\alpha$ on $\mathcal{S}(\mathbf{C})$ by

$$\mathscr{H}_\alpha \phi(z) := \frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \int_{|t| \geq \varepsilon} \frac{\phi(z - e^{i\alpha} t)}{t} \, \mathrm{d}t,$$

where the integral is over the real coordinate t. $\diamondsuit$

The angular Hilbert transforms are seen to be well-defined in much the same way as this has been done for the Hilbert transform.

2.37 Lemma. *Let $\alpha \in \mathbf{R}$ and $p \in]1, \infty[$. Then $\mathscr{H}_\alpha$ has an extension in $\mathcal{L}(L^p(\mathbf{C}))$. If we again denote this extension by $\mathscr{H}_\alpha$, then we have $\|\mathscr{H}_\alpha\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq \|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))}$.*

Proof. First we observe that

$$\mathscr{H}_\alpha = r^*_{-\alpha} \circ \mathscr{H}_0 \circ r^*_\alpha.$$

Let $\phi \in \mathcal{S}(\mathbf{C})$. Assuming we have proven the result in the case that $\alpha = 0$, we find

$$\|\mathscr{H}_\alpha \phi\|_p \leq \|r^*_{-\alpha}\|_{\mathcal{L}(L^p(\mathbf{C}))} \|\mathscr{H}_0\|_{\mathcal{L}(L^p(\mathbf{C}))} \|r^*_\alpha\|_{\mathcal{L}(L^p(\mathbf{C}))} \|\phi\|_p \leq \|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))} \|\phi\|_p.$$

This shows that it suffices to consider the case where $\alpha = 0$.

For $y \in \mathbf{R}$ we define $\psi_y \in \mathcal{S}(\mathbf{R})$ by $\psi_y(x) := \phi(x + iy)$. Then it follows from the definition of $\mathscr{H}_0$ that $\mathscr{H}\psi_y(x) = \mathscr{H}_0\phi(x + iy)$. Hence,

$$\|\mathscr{H}_0 \phi\|_p^p \leq \int_{\mathbf{R}} \|\mathscr{H}\|^p_{\mathcal{L}(L^p(\mathbf{R}))} \int_{\mathbf{R}} |\psi_y(x)|^p \, \mathrm{d}x \, \mathrm{d}y = \|\mathscr{H}\|^p_{\mathcal{L}(L^p(\mathbf{R}))} \|\phi\|_p^p,$$

by Fubini's Theorem. The assertion follows. $\square$

Proof of Theorem 2.34. We will prove the result for $\mathscr{R}_1$. The corresponding result for $\mathscr{R}_2$ is completely analogous. Let $\varepsilon \in \mathbf{R}_+$, $w \in \mathbf{C}$, and $\phi \in C_c^\infty(\mathbf{C})$. Then, by subsequently changing to polar coordinates and substituting α for $\alpha + \pi$, we obtain

$$\int_{|z| \geq \varepsilon} \frac{x}{|z|^3} \phi(w - z) \, \mathrm{d}z = \int_0^{2\pi} \cos\alpha \int_\varepsilon^\infty \frac{\phi(w - re^{i\alpha})}{r} \, \mathrm{d}r \, \mathrm{d}\alpha \tag{2.51}$$

$$= -\int_0^{2\pi} \cos\alpha \int_\varepsilon^\infty \frac{\phi(w + re^{i\alpha})}{r} \, \mathrm{d}r \, \mathrm{d}\alpha \tag{2.52}$$

Adding (2.51) and (2.52) and dividing by 2 yields

$$\int_{|z| \geq \varepsilon} \frac{x}{|z|^3} \phi(w - z) \, \mathrm{d}z = \frac{1}{2} \int_0^{2\pi} \cos\alpha \int_\varepsilon^\infty \frac{\phi(w - re^{i\alpha}) - \phi(w + re^{i\alpha})}{r} \, \mathrm{d}r \, \mathrm{d}\alpha. \tag{2.53}$$

We now wish to justify using Lebesgue's Dominated Convergence Theorem to conclude that

$$\mathscr{R}_1 \phi(w) = \frac{1}{4} \int_0^{2\pi} \cos\alpha \mathscr{H}_\alpha \phi(w) \, \mathrm{d}\alpha. \tag{2.54}$$

For this, note that for all $\alpha \in [0, 2\pi]$ we have

$$\left| \frac{\phi(w - re^{i\alpha}) - \phi(w + re^{i\alpha})}{r} \right| \leq 2^{\frac{3}{2}} \nu_{0,1}(\phi),$$

by applying the Mean Value Theorem to the function $r \mapsto \phi(w - re^{i\alpha})$. Then, by compactness of the support of ϕ, we can find some $R > \varepsilon$ so that

$$\left| \int_\varepsilon^\infty \frac{\phi(w - re^{i\alpha}) - \phi(w + re^{i\alpha})}{r} \, \mathrm{d}r \right| \leq \int_\varepsilon^R \left| \frac{\phi(w - re^{i\alpha}) - \phi(w + re^{i\alpha})}{r} \right| \mathrm{d}r \leq 2^{\frac{3}{2}} R \nu_{0,1}(\phi).$$

This justifies letting $\varepsilon \downarrow 0$ in (2.53) to conclude (2.54).

If $p \in]1, \infty[$, then we can use Minkowski's integral inequality and Lemma 2.37 to conclude that

$$\|\mathscr{R}_1 \phi\|_p \leq \frac{1}{4} \int_0^{2\pi} |\cos \alpha| \, \mathrm{d}\alpha \|\mathscr{H}_\alpha \phi\|_p \leq \|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))} \|\phi\|_p,$$

since

$$\int_0^{2\pi} |\cos \alpha| \, \mathrm{d}\alpha = 4.$$

The assertion now follows from Lemma 1.1. □

In order to prove Theorem 2.35, we first consider another tempered distribution. Note that the function $|z|^{-1}$ is locally integrable in $\mathbf{C}$, since

$$\int_{|z| \leq 1} \frac{1}{|z|} \, \mathrm{d}z = 2\pi \int_0^1 r \frac{1}{r} \, \mathrm{d}r = 2\pi < \infty.$$

Thus, $|z|^{-1}$ defines a distribution. However, it actually defines a tempered distribution, since

$$\begin{aligned} \int_{\mathbf{C}} \frac{|\phi(z)|}{|z|} \, \mathrm{d}z &= \int_{|z| \leq 1} \frac{|\phi(z)|}{|z|} \, \mathrm{d}z + \int_{|z| \geq 1} \frac{|\phi(z)|}{|z|} \, \mathrm{d}z \\ &\leq 2\pi \nu_{0,0}(\phi) + \nu_{2,0}(\phi) \int_{|z| \geq 1} \frac{1}{|z|^3} \, \mathrm{d}z \\ &= 2\pi \nu_{0,0}(\phi) + 2\pi \nu_{2,0}(\phi) \end{aligned} \tag{2.55}$$

for all $\phi \in \mathcal{S}(\mathbf{R})$. This connects to the Riesz transforms as follows:

2.38 Lemma. *We have*

$$-\partial_x |z|^{-1} = \mathrm{PV} \frac{x}{|z|^3}, \quad -\partial_y |z|^{-1} = \mathrm{PV} \frac{y}{|z|^3},$$

where the derivatives are taken in $\mathcal{S}'(\mathbf{C})$.

Proof. Let $\varepsilon \in \mathbf{R}_+$ and $\phi \in \mathcal{S}(\mathbf{C})$. Denote by σ_ε the standard surface measure on the circle $S(\varepsilon) = \{\omega = \omega_x + i\omega_y \in \mathbf{C} \mid |\omega| = \varepsilon\}$. As the outward unit normal vector ν to $S(\varepsilon)$, seen as the boundary of $\{z \in \mathbf{C} \mid |z| \geq \varepsilon\}$, is given by $\nu(z) = -z/\varepsilon$, we have, using partial integration,

$$\int_{|z|\geq\varepsilon} \partial_x\phi(z)|z|^{-1}\,\mathrm{d}z = \int_{|z|\geq\varepsilon} \frac{x}{|z|^3}\phi(z)\,\mathrm{d}z - \int_{S(\varepsilon)} \frac{\omega_x}{\varepsilon^2}\phi(\omega)\,\mathrm{d}\sigma_\varepsilon(\omega), \tag{2.56}$$

and similarly for y instead of x. Since

$$\lim_{\varepsilon\downarrow 0}\int_{S(\varepsilon)} \frac{\omega_x}{\varepsilon^2}\phi(\omega)\,\mathrm{d}\sigma_\varepsilon(\omega) = \lim_{\varepsilon\downarrow 0}\int_0^{2\pi} \phi(\varepsilon e^{it})\cos(t)\,\mathrm{d}t = \phi(0)\int_0^{2\pi}\cos(t)\,\mathrm{d}t = 0,$$

and similarly for y where cos is replaced by sin, we may take the limit as $\varepsilon \downarrow 0$ in (2.56) to conclude that

$$\langle -\partial_x|z|^{-1}, \phi\rangle = \langle |z|^{-1}, \partial_x\phi\rangle = \lim_{\varepsilon\downarrow 0}\int_{|z|\geq\varepsilon} \frac{x}{|z|^3}\phi(z)\,\mathrm{d}z = \left\langle \mathrm{PV}\,\frac{x}{|z|^3}, \phi\right\rangle,$$

and similarly for y. This proves the desired result. ☐

In view of Lemma 2.38, the vital step for proving Theorem 2.35 is the computation of the Fourier transform of $|z|^{-1}$. As it turns out, this is actually a fixed point of the Fourier transform. To show this, we will use the following result:

2.39 Lemma. *Suppose $f \in C^\infty(\mathbf{C}\backslash\{0\})$ satisfies $f(tz) = t^{-1}f(z)$ and $f(e^{i\alpha}z) = f(z)$ for all $z \in \mathbf{C}\backslash\{0\}$, $t \in \mathbf{R}_+$, and $\alpha \in \mathbf{R}$. Then there exists a constant $c \in \mathbf{C}$ so that $f = c|z|^{-1}$.*

Proof. Write $z \in \mathbf{C}\backslash\{0\}$ in polar coordinates as $z = re^{i\alpha}$ for $r \in \mathbf{R}_+$ and $\alpha \in \mathbf{R}$. Then

$$f(z) = f(re^{i\alpha}) = f(e^{i\alpha})r^{-1} = f(1)|z|^{-1}.$$

The assertion then follows with $c = f(1)$. ☐

2.40 Lemma. *The Fourier transform of $|z|^{-1}$ is given by $|\zeta|^{-1}$.*

Proof. Since $|z|^{-1}$ is invariant under rotations and homogeneous of degree -1, see Example B.18, it follows from Corollary B.33 that $\mathscr{F}|z|^{-1}$ is also invariant under rotations and homogeneous of degree $-1 = -2-(-1)$. Moreover, since $|z|^{-1}$ is smooth in $\mathbf{C}\backslash\{0\}$, we conclude from Theorem B.38 that $\mathscr{F}|z|^{-1}$ is given by a smooth function f in $\mathbf{C}\backslash\{0\}$. From this it follows that $f(tz) = t^{-1}f(z)$ for all $t \in \mathbf{R}_+$ and $z \in \mathbf{C}\backslash\{0\}$. Hence, f satisfies the conditions from Lemma 2.39 and is thus of the form $f = c|\zeta|^{-1}$ for some $c \in \mathbf{C}$. Thus, f defines a tempered distribution in its own right.

Next, note that $u := \mathscr{F}|z|^{-1} - f \in \mathcal{S}'(\mathbf{C})$ is supported in the origin. This implies, by Theorem B.12, that u is of the form

$$u = \sum_{|\alpha|\leq k} c_\alpha \partial^\alpha\delta,$$

where $k \in \mathbf{Z}_{\geq 0}$ and $c_\alpha \in \mathbf{C}$ for all multi-indices α with $|\alpha| \leq k$. As is shown in Example B.18, the distributions $\partial^\alpha\delta$ are homogeneous of degree $-2-|\alpha|$ for every multi-index α, while u is homogeneous of degree -1. Hence, for every $t \in \mathbf{R}_+$ we have

$$t^{-1}u = d_t u = \sum_{|\alpha|\leq k} c_\alpha d_t(\partial^\alpha\delta) = \sum_{|\alpha|\leq k} t^{-2-|\alpha|}c_\alpha\partial^\alpha\delta,$$

which implies that for all $\phi \in \mathcal{S}(\mathbf{C})$ we have

$$u(\phi) = \sum_{|\alpha|\le k} t^{-1-|\alpha|} c_\alpha \partial^\alpha \delta(\phi) \to 0 \quad \text{as} \quad t \to \infty$$

so that $u = 0$. We conclude that $\mathscr{F}|z|^{-1} = c|\zeta|^{-1}$.

To compute c, we recall that if $\gamma := e^{-\pi|z|^2} \in \mathcal{S}(\mathbf{C})$, then $\mathscr{F}\gamma = \gamma$, see Example B.29. Hence,

$$c\langle|\zeta|^{-1}, \gamma\rangle = \langle\mathscr{F}|z|^{-1}, \gamma\rangle = \langle|z|^{-1}, \gamma\rangle.$$

Since $\langle|z|^{-1}, \gamma\rangle > 0$, we conclude that $c = 1$. The assertion follows. □

Proof of Theorem 2.35. By density, it suffices to show the result for functions in $\mathcal{S}(\mathbf{C})$. Let $\phi \in \mathcal{S}(\mathbf{C})$. From Lemma 2.38 and Lemma 2.40 it follows that

$$\mathscr{F}(\mathscr{R}_1\phi) = -\frac{1}{2\pi}\mathscr{F}(\partial_x|z|^{-1})\mathscr{F}\phi = -i\xi\mathscr{F}|z|^{-1}\mathscr{F}\phi = -i\frac{\xi}{|\zeta|}\mathscr{F}\phi.$$

The proof for $\mathscr{R}_2$ is analogous. The assertion follows. □

As an analogue to Corollary 2.22 we also show the following:

2.41 Corollary. *Let $p \in]1, \infty[$ and $f \in L^p(\mathbf{C})$. Then*

$$\mathscr{R}_1^2 f + \mathscr{R}_2^2 f = -f.$$

Proof. Since

$$\left(-i\frac{x}{|z|}\right)^2 + \left(-i\frac{y}{|z|}\right)^2 = -1,$$

the result follows by taking the Fourier transform. □

2.42 Definition. Let $p \in]1, \infty[$. Denote the extensions of $\mathscr{R}_1$ and $\mathscr{R}_2$ to $L^p(\mathbf{C})$ again by $\mathscr{R}_1$ and $\mathscr{R}_2$. The *Beurling-Ahlfors transform* $\mathscr{B}_p$ in $L^p(\mathbf{C})$ is defined as the operator $\mathscr{B}_p := (i\mathscr{R}_1 + \mathscr{R}_2)^2 \in \mathcal{L}(L^p(\mathbf{C}))$. ♢

We note that, per definition, all the operators $(\mathscr{B}_p)_{p\in]1,\infty[}$ coincide on $\mathcal{S}(\mathbf{C})$ and in particular on $C_c^\infty(\mathbf{C})$. Recalling Lemma 2.5, we could denote all these operators simply by $\mathscr{B}$. We opt to not do this at this point for the sake of clarity and to emphasize the involved subtleties.

2.43 Proposition. *For all $f \in L^2(\mathbf{C})$ we have*

$$\mathscr{F}(\mathscr{B}_2 f) = M_{\frac{\bar\zeta}{\zeta}}\mathscr{F}f.$$

Proof. We have

$$\mathscr{F}(\mathscr{B}_2 f) = (iM_{-i\frac{\xi}{|\zeta|}} + M_{-i\frac{\eta}{|\zeta|}})^2\mathscr{F}f = M^2_{\frac{\xi - i\eta}{|\zeta|}}\mathscr{F}f = M_{\frac{\bar\zeta}{\zeta}}\mathscr{F}f,$$

as desired. □

This proposition can now be used to show that the Beurling-Ahlfors transform interchanges the Wirtinger derivatives. Before we do this, we note, recalling Remark 1.3, that it is straightforward to check that the conjugate transpose operator $\mathscr{B}_2^\dagger$ satisfies

$$\mathscr{B}_2^\dagger = (-i\mathscr{R}_1 + \mathscr{R}_2)^2 = \mathscr{F}^{-1} M_{\frac{z}{\bar{z}}} \mathscr{F}.$$

This first equality also gives a description of $\mathscr{B}_p^\dagger$ for all $p \in]1, \infty[$, while the second equality implies the following:

2.44 Lemma. *Let $p \in]1, \infty[$. Then $\mathscr{B}_p$ is an isomorphism of $L^p(\mathbf{C})$ with inverse $\mathscr{B}_{p'}^\dagger$. In particular, $\mathscr{B}_2$ is a unitary isomorphism of $L^2(\mathbf{C})$.*

Proof. Since

$$M_{\frac{\bar{z}}{z}} M_{\frac{z}{\bar{z}}} = M_{\frac{z}{\bar{z}}} M_{\frac{\bar{z}}{z}} = \mathrm{id}_{L^2(\mathbf{C})},$$

it follows from taking the Fourier transform that $\mathscr{B}_2 \mathscr{B}_2^\dagger$ and $\mathscr{B}_2^\dagger \mathscr{B}_2$ coincide with the identity mapping on $C_c^\infty(\mathbf{C})$. The assertion for $p = 2$ follows by density. By Lemma 2.5 we find that for all $\phi \in C_c^\infty(\mathbf{C})$ we have $\mathscr{B}_2^\dagger \phi = \mathscr{B}_{p'}^\dagger \phi \in L^2(\mathbf{C}) \cap L^{p'}(\mathbf{C})$. Hence, another application of Lemma 2.5 implies that

$$\mathscr{B}_p \mathscr{B}_{p'}^\dagger \phi = \mathscr{B}_2 \mathscr{B}_{p'}^\dagger \phi = \mathscr{B}_2 \mathscr{B}_2^\dagger \phi = \phi$$

for all $\phi \in C_c^\infty(\mathbf{C})$. Thus, $\mathscr{B}_p \mathscr{B}_{p'}^\dagger$ coincides with the identity mapping on $C_c^\infty(\mathbf{C})$. Analogously we find that $\mathscr{B}_{p'}^\dagger \mathscr{B}_p$ also coincides with the identity mapping on $C_c^\infty(\mathbf{C})$. The assertion now follows by density. □

We wish to establish a lower bound on the operator norms of the Beurling-Ahlfors transform.

2.45 Proposition. *Let $p \in]1, \infty[$. Then we have $\|\mathscr{B}_p\|_{\mathcal{L}(L^p(\mathbf{C}))} \geq p^* - 1$.*

This proposition will be proved by computing the Beurling-Ahlfors transform of specific functions with favorable partial derivatives. Such examples were first discovered by Lehto, see [Le].

For $p \in [1, \infty[$, we define the *homogeneous Sobolev space*

$$\mathring{W}^{1,p}(\mathbf{C}) := \{f \in \mathcal{D}'(\mathbf{C}) \mid \partial_z f, \partial_{\bar{z}} f \in L^p(\mathbf{C})\},$$

topologized by the seminorm

$$\|f\|_{\mathring{W}^{1,p}(\mathbf{C})} := \|\partial_z f\|_p + \|\partial_{\bar{z}} f\|_p.$$

2.46 Proposition. *Let $p \in]1, \infty[$. For any $f \in \mathring{W}^{1,p}(\mathbf{C})$ we have*

$$\mathscr{B}_p(\partial_{\bar{z}} f) = \partial_z f, \quad \mathscr{B}_p^\dagger(\partial_z f) = \partial_{\bar{z}} f.$$

As one might expect, the proof uses a density argument.

2.47 Lemma. *For all $p \in [1, \infty[$, the space $C_c^\infty(\mathbf{C})$ is dense in $\mathring{W}^{1,p}(\mathbf{C})$.*

Proof. This proof is nearly identical to the proof of Theorem 1.5 and will therefore be omitted. □

Proof of Proposition 2.46. The equality for $\mathscr{B}_p^\dagger$ follows from the equality for $\mathscr{B}_p$ by Lemma 2.44. Let $\phi \in C_c^\infty(\mathbf{C})$. Then

$$\mathscr{F}(\mathscr{B}_2(\partial_{\overline{z}}\phi)) = \frac{\overline{z}}{z}\mathscr{F}(\partial_{\overline{z}}\phi) = \pi i \overline{z}\mathscr{F}\phi = \mathscr{F}(\partial_z \phi),$$

hence $\mathscr{B}_p(\partial_{\overline{z}}\phi) = \mathscr{B}_2(\partial_{\overline{z}}\phi) = \partial_z\phi$. The result now follows from density of $C_c^\infty(\mathbf{C})$ in $\mathring{W}^{1,p}(\mathbf{C})$. $\square$

As an analogue to Proposition 2.6, we have the following:

2.48 Proposition. *Let $p \in]1,\infty[$. Then $\mathscr{B}_p^* = \mathscr{B}_{p'}$.*

Proof. We note that

$$\mathscr{B}_2^* = \mathscr{F} M_{\frac{\overline{z}}{z}} \mathscr{F}^{-1} = \mathscr{F}^{-1} M_{\frac{\overline{-z}}{-z}} \mathscr{F} = \mathscr{B}_2,$$

implying that

$$\langle \mathscr{B}_2 f, g\rangle = \langle f, \mathscr{B}_2^* g\rangle = \langle f, \mathscr{B}_2 g\rangle$$

for all $f, g \in L^2(\mathbf{C})$. By Lemma 2.5, this means that

$$\langle \mathscr{B}_p f, g\rangle = \langle f, \mathscr{B}_{p'} g\rangle$$

for all $f \in L^2(\mathbf{C}) \cap L^p(\mathbf{C})$ and $g \in L^2(\mathbf{C}) \cap L^{p'}(\mathbf{C})$. The assertion now follows from density of $L^2(\mathbf{C}) \cap L^p(\mathbf{C})$ and $L^2(\mathbf{C}) \cap L^{p'}(\mathbf{C})$ in $L^p(\mathbf{C})$ and $L^{p'}(\mathbf{C})$ respectively. $\square$

Finally, before we prove the lower bound on the operator norms of the Beurling-Ahlfors transform, we consider certain functions known as *radial stretchings*.

2.49 Example. Suppose a function $f : \mathbf{C} \to \mathbf{C}$ is of the form

$$f(z) = \begin{cases} \frac{z}{|z|}\rho(|z|) & \text{if } |z| > 0 \\ 0 & \text{if } z = 0, \end{cases}$$

where $\rho : \mathbf{R}_+ \to \mathbf{R}_+$ is a strictly increasing continuously differentiable function that extends continuously to $\mathbf{R}_{\geq 0}$ by $\rho(0) = 0$. Then f is continuous and continuously differentiable outside of the origin. A direct computation shows that $\partial_z|z| = \overline{z}/(2|z|)$ and thus $\partial_{\overline{z}}|z| = \overline{\partial_z|z|} = z/(2|z|)$. Using Proposition C.1, this implies that

$$\partial_z \frac{z}{|z|} = \partial_z \frac{|z|}{\overline{z}} = \frac{1}{2|z|}, \quad \partial_{\overline{z}} \frac{z}{|z|} = -\frac{z}{|z|^2}\frac{z}{2|z|} = -\frac{z}{\overline{z}}\frac{1}{2|z|}.$$

Hence,

$$\begin{aligned} \partial_z f(z) &= \frac{\rho(|z|)}{2|z|} + \frac{z}{|z|}\rho'(|z|)\frac{\overline{z}}{2|z|} = \frac{1}{2}\left(\rho'(|z|) + \frac{\rho(|z|)}{|z|}\right) \\ \partial_{\overline{z}} f(z) &= -\frac{z}{\overline{z}}\frac{\rho(|z|)}{2|z|} + \frac{z}{|z|}\rho'(|z|)\frac{z}{2|z|} = \frac{z}{2\overline{z}}\left(\rho'(|z|) - \frac{\rho(|z|)}{|z|}\right). \end{aligned}$$

It follows that f satisfies

$$\partial_{\overline{z}} f(z) = \mu(z) \partial_z f(z)$$

for a.e. $z \in \mathbf{C}$, where $\mu \in L^\infty(\mathbf{C})$ is defined by

$$\mu(z) = \frac{z}{\overline{z}} \frac{|z|\rho'(|z|) - \rho(|z|)}{|z|\rho'(|z|) + \rho(|z|)}$$

for $z \in \mathbf{C}\backslash\{0\}$. ♢

Proof of Proposition 2.45. It follows from Lemma 2.44 that $\|\mathscr{B}_2\|_{\mathcal{L}(L^2(\mathbf{C}))} = 1$. Thus, since $p^* - 1 = (p')^* - 1$, by Proposition 2.48 it is now sufficient to consider the cases $p \in]2, \infty[$ to establish the result.

Let $p \in]2, \infty[$, $\gamma \in]0, 1[$, and define $f : \mathbf{C} \to \mathbf{C}$ by

$$f(z) := \begin{cases} 0 & \text{if } z = 0 \\ z|z|^{-\frac{2\gamma}{p}} & \text{if } 0 < |z| < 1 \\ \dfrac{1}{\overline{z}} & \text{if } |z| \geq 1. \end{cases}$$

Since $1 - 2\gamma/p = (p - 2\gamma)/p \in \mathbf{R}_+$ and $z\overline{z} = 1$ for $|z| = 1$, this function is continuous in $\mathbf{C}$, hence, lies in $L^1_{loc}(\mathbf{C})$. This means that f defines a distribution in $\mathbf{C}$. Setting $\rho(t) := t^{(p-2\gamma)/p}$ for $t \in \mathbf{R}_+$, we note that $z|z|^{-\frac{2\gamma}{p}} = z/|z|\rho(|z|)$. Thus, by Example 2.49, we note that f is continuously differentiable for $0 < |z| < 1$ and $|z| > 1$ with

$$\partial_z f(z) = \begin{cases} \dfrac{1}{2}\left(|z|^{-\frac{2\gamma}{p}} + \dfrac{p - 2\gamma}{p}|z|^{-\frac{2\gamma}{p}}\right) = \dfrac{p-\gamma}{p}|z|^{-\frac{2\gamma}{p}} & \text{if } 0 < |z| < 1 \\ 0 & \text{if } |z| > 1, \end{cases}$$

and

$$\partial_{\overline{z}} f(z) = \begin{cases} \dfrac{z}{2\overline{z}}\left(-|z|^{-\frac{2\gamma}{p}} + \dfrac{p - 2\gamma}{p}|z|^{-\frac{2\gamma}{p}}\right) = -\dfrac{\gamma}{p}\dfrac{z}{\overline{z}}|z|^{-\frac{2\gamma}{p}} & \text{if } 0 < |z| < 1 \\ -\dfrac{1}{\overline{z}^2} & \text{if } |z| > 1. \end{cases}$$

These a.e. defined functions are locally integrable in their own right and hence define distributions. We claim that they are the distributional derivatives of f. Indeed, let S^1 denote the unit circle in $\mathbf{C}$ and εS^1 the circle around 0 of radius ε for $\varepsilon \in \mathbf{R}_+$ and let $\phi \in C_c^\infty(\mathbf{C})$. Then, since $\partial_z(f\phi) = (\partial_z f)\phi + f(\partial_z \phi)$ where $\partial_z f$ is defined, we note that by Green's Integral Theorem we have

$$\begin{aligned}
-\int_{\mathbf{C}} f(z)\partial_z\phi(z)\,\mathrm{d}z &= -\int_{0<|z|<1} f(z)\partial_z\phi(z)\,\mathrm{d}z - \int_{|z|>1} f(z)\partial_z\phi(z)\,\mathrm{d}z \\
&= \int_{|z|<1} \partial_z f(z)\phi(z)\,\mathrm{d}z - \lim_{\varepsilon\downarrow 0}\int_{\varepsilon<|z|<1} \partial_z(\phi f)(z)\,\mathrm{d}z + \int_{|z|>1} \partial_z f(z)\phi(z)\,\mathrm{d}z - \int_{|z|>1} \partial_z(\phi f)(z)\,\mathrm{d}z \\
&= \int_{\mathbf{C}} \partial_z f(z)\phi(z)\,\mathrm{d}z + \lim_{\varepsilon\downarrow 0}\oint_{\varepsilon S^1} f(z)\phi(z)\,\mathrm{d}\overline{z} - \oint_{S^1} f(z)\phi(z)\,\mathrm{d}\overline{z} + \oint_{S^1} f(z)\phi(z)\,\mathrm{d}\overline{z} \\
&= \int_{\mathbf{C}} \partial_z f(z)\phi(z)\,\mathrm{d}z + \lim_{\varepsilon\downarrow 0}\oint_{\varepsilon S^1} f(z)\phi(z)\,\mathrm{d}\overline{z},
\end{aligned}$$

where the orientation of the circles is counterclockwise and

$$\lim_{\varepsilon\downarrow 0}\oint_{\varepsilon S^1} f(z)\phi(z)\,\mathrm{d}\overline{z} = \lim_{\varepsilon\downarrow 0} -i\varepsilon\int_0^{2\pi} f(\varepsilon e^{it})\phi(\varepsilon e^{it})e^{-it}\,\mathrm{d}t = 0$$

by continuity of f at 0. The computation for $\partial_{\overline{z}} f$ is similar, which proves the claim.

Thus, using polar coordinates we may compute

$$\|\partial_z f\|_p^p = 2\pi\left(\frac{p-\gamma}{p}\right)^p\int_0^1 r^{1-2\gamma}\,\mathrm{d}r = \frac{\pi}{1-\gamma}\left(\frac{p-\gamma}{p}\right)^p,$$

and

$$\|\partial_{\overline{z}} f\|_p^p = \frac{\pi}{1-\gamma}\left(\frac{\gamma}{p}\right)^p + 2\pi\int_1^\infty r^{1-2p}\,\mathrm{d}r = \frac{\pi}{1-\gamma}\left(\frac{\gamma}{p}\right)^p + \frac{\pi}{p-1}.$$

In particular, we have establishes that $f \in \mathring{W}^{1,p}(\mathbf{C})$ (and actually, we have $f \in W^{1,p}(\mathbf{C})$). Hence, by Proposition 2.46,

$$\frac{\|\mathscr{B}_p(\partial_{\overline{z}} f)\|_p}{\|\partial_{\overline{z}} f\|_p} = \frac{\|\partial_z f\|_p}{\|\partial_{\overline{z}} f\|_p} = \left(\frac{\left(\frac{p-\gamma}{p}\right)^p}{\left(\frac{\gamma}{p}\right)^p + \frac{1-\gamma}{p-1}}\right)^{\frac{1}{p}},$$

so that we may now conclude that

$$\|\mathscr{B}_p\|_{\mathcal{L}(L^p(\mathbf{C}))} \geq \left(\frac{\left(\frac{p-\gamma}{p}\right)^p}{\left(\frac{\gamma}{p}\right)^p + \frac{1-\gamma}{p-1}}\right)^{\frac{1}{p}} \tag{2.57}$$

for all $\gamma \in]0,1[$. The right-hand side of (2.57) tends to $p-1 = p^*-1$ as $\gamma \uparrow 1$. The assertion follows. □

We can also establish that the operator norms of the Beurling-Ahlfors transform are determined by the interchange of the Wirtinger derivatives.

2.50 Proposition. *Let $p \in]1,\infty[$. Suppose V is a vector space so that $C_c^\infty(\mathbf{C}) \subseteq V \subseteq \mathring{W}^{1,p}(\mathbf{C})$. Then*

$$\begin{aligned}\|\mathscr{B}_p\|_{\mathcal{L}(L^p(\mathbf{C}))} &= \inf\{c \in \mathbf{R}_+ \mid \|\partial_z f\|_p \leq c\|\partial_{\overline{z}} f\|_p \text{ for all } f \in V\}\\ &= \inf\{c \in \mathbf{R}_+ \mid \|\partial_{\overline{z}} f\|_p \leq c\|\partial_z f\|_p \text{ for all } f \in V\}.\end{aligned}$$

This uses density of $\partial_{\overline{z}}(C_c^\infty(\mathbf{C})) = \{\partial_{\overline{z}}\phi \mid \phi \in C_c^\infty(\mathbf{C})\}$ in $L^p(\mathbf{C})$ for $p \in]1,\infty[$. For this we will use the following extension of Liouville's Theorem:

2.51 Lemma. *Let $p \in [1,\infty]$. If $f \in L^p(\mathbf{C})$ satisfies $\partial_{\overline{z}} f = 0$ in $\mathcal{D}'(\mathbf{C})$, then f is constant if $p = \infty$ and $f = 0$ if $p \in [1,\infty[$.*

Proof. Note that $\pi i\zeta\mathscr{F} f = \mathscr{F}(\partial_{\overline{z}} f) = 0$, which implies that $\operatorname{supp}\mathscr{F} f \subseteq \{0\}$. Hence, $\mathscr{F} f$ must be a linear combination of derivatives of the Dirac delta distribution. But this means that f itself is a polynomial. The only bounded polynomials in $\mathbf{C}$ are the constant ones, while the only p-integrable polynomial in $\mathbf{C}$ for $p \in [1,\infty[$ is the zero polynomial. The assertion follows. □

2.52 Lemma. *Let $p \in]1,\infty[$. Then $\partial_{\overline{z}}(C_c^\infty(\mathbf{C}))$ is dense in $L^p(\mathbf{C})$.*

Proof. The asserted density is equivalent to showing that the annihilator of $\partial_{\overline{z}}(C_c^\infty(\mathbf{C}))$ consists of only the 0 function. This means that we have to show that if

$$\int_{\mathbf{C}} f(z)\partial_{\overline{z}}\phi(z)\,\mathrm{d}z = 0$$

for all $\phi \in C_c^\infty(\mathbf{C})$ for some $f \in L^{p'}(\mathbf{C})$, then $f = 0$. But this is precisely what Lemma 2.51 asserts. This proves the result. □

We note that the closure of $\partial_{\overline{z}}(C_c^\infty(\mathbf{C}))$ in $L^1(\mathbf{C})$ is given by the space of those integrable functions that have mean value zero. In particular, $\partial_{\overline{z}}(C_c^\infty(\mathbf{C}))$ is not dense in $L^1(\mathbf{C})$ or $L^\infty(\mathbf{C})$.

Proof of Proposition 2.50. In view of Proposition 2.46, the second equality follows in the same way as the first one by replacing $\mathscr{B}_p$ by $\mathscr{B}_{p'}^\dagger$ and by noting that $\|\mathscr{B}_p\|_{\mathcal{L}(L^p(\mathbf{C}))} = \|\mathscr{B}_{p'}^\dagger\|_{\mathcal{L}(L^p(\mathbf{C}))}$. Thus, we will only show the first equality.

For each vector space $C_c^\infty(\mathbf{C}) \subseteq W \subseteq \mathring{W}^{1,p}(\mathbf{C})$ we set

$$c(W) := \inf\{c \in \mathbf{R}_+ \mid \|\partial_z f\|_p \leq c\|\partial_{\overline{z}} f\|_p \text{ for all } f \in W\}.$$

Then, by Proposition 2.46, the chain of inequalities

$$c(C_c^\infty(\mathbf{C})) \leq c(V) \leq c(\mathring{W}^{1,p}(\mathbf{C})) \leq \|\mathscr{B}_p\|_{\mathcal{L}(L^p(\mathbf{C}))}$$

is clear. Hence, it suffices to show that $\|\mathscr{B}_p\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq c(C_c^\infty(\mathbf{C}))$.

Note that for all $\phi \in C_c^\infty(\mathbf{C})$ we have

$$\|\mathscr{B}_p(\partial_{\overline{z}}\phi)\|_p = \|\partial_z\phi\|_p \leq c(C_c^\infty(\mathbf{C}))\|\partial_{\overline{z}}\phi\|_p.$$

But then it follows from Lemma 2.52 that

$$\|\mathscr{B}_p f\|_p \leq c(C_c^\infty(\mathbf{C}))\|f\|_p$$

for all $f \in L^p(\mathbf{C})$. This proves that $\|\mathscr{B}_p\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq c(C_c^\infty(\mathbf{C}))$, as desired. □

Finally, we will give an alternative method of defining the Beurling-Ahlfors transform, as a complex analytic tool.

Consider the locally integrable function $E := (\pi z)^{-1}$. Then we note that the estimate (2.55) shows us that $E \in \mathcal{S}'(\mathbf{C})$.

2.53 Definition. We define the *Cauchy transform* $\mathscr{C} : \mathcal{S}(\mathbf{C}) \to \mathcal{O}_M(\mathbf{C})$ as the convolution operator

$$\mathscr{C}\phi := E * \phi.$$

◊

We denote by δ the Dirac delta distribution at the origin. By Proposition C.12, the tempered distribution E satisfies $\partial_{\overline{z}} E = \delta$ in $\mathcal{D}'(\mathbf{C})$. By Proposition B.43 this implies that

$$\partial_{\overline{z}} \mathscr{C}(\phi) = \partial_{\overline{z}} E * \phi = \phi$$

for all $\phi \in \mathcal{S}(\mathbf{C})$.

2.54 Definition. We define the Beurling-Ahlfors transform $\mathscr{B} : \mathcal{S}(\mathbf{C}) \to \mathcal{O}_M(\mathbf{C})$ as the convolution operator

$$\mathscr{B}\phi := \partial_z(\mathscr{C}\phi) = \partial_z E * \phi = E * \partial_z \phi.$$

$\diamondsuit$

It follows from this definition that we have the relations

$$\partial_{\overline{z}}(\mathscr{C}\phi) = \mathscr{C}(\partial_{\overline{z}}\phi) = \phi, \quad \partial_z(\mathscr{C}\phi) = \mathscr{C}(\partial_z \phi) = \mathscr{B}\phi \quad \text{for all } \phi \in \mathcal{S}(\mathbf{C}). \tag{2.58}$$

In particular, this means that

$$\mathscr{B}(\partial_{\overline{z}}\phi) = \partial_z(\mathscr{C}(\partial_{\overline{z}}\phi)) = \partial_{\overline{z}}(\mathscr{C}(\partial_z \phi)) = \partial_z \phi.$$

for all $\phi \in \mathcal{S}(\mathbf{C})$.

2.55 Proposition. *We have*

$$\mathscr{B}\phi(w) := -\frac{1}{\pi} \operatorname{PV} \frac{1}{z^2} * \phi(w) = -\frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \int_{|w-z| \geq \varepsilon} \frac{\phi(z)}{(w-z)^2} \, \mathrm{d}z.$$

for all $\phi \in \mathcal{S}(\mathbf{C})$ *and* $w \in \mathbf{C}$.

Proof. First we must check that $\operatorname{PV} 1/z^2$ is a well-defined tempered distribution. Then, it suffices to show that we have the equality of distributions

$$\partial_z \frac{1}{z} = -\operatorname{PV} \frac{1}{z^2} \tag{2.59}$$

to conclude the proof.

Let $\phi \in \mathcal{S}(\mathbf{C})$ and let $\varepsilon \in \mathbf{R}_+$ so that $\varepsilon < 1$. By (C.3) in Example C.9 we find that the integral of $1/z^2$ over the annulus $\varepsilon \leq |z| \leq 1$ vanishes. Hence,

$$\left| \lim_{\varepsilon \downarrow 0} \int_{\varepsilon \leq |z| \leq 1} \frac{\phi(z)}{z^2} \, \mathrm{d}z \right| = \left| \lim_{\varepsilon \downarrow 0} \int_{\varepsilon \leq |z| \leq 1} \frac{\phi(z) - \phi(0)}{z^2} \, \mathrm{d}z \right| \leq \nu_{0,1}(\phi) \int_{|z| \leq 1} \frac{1}{|z|} \, \mathrm{d}z = 2\pi \nu_{0,1}(\phi).$$

Since

$$\int_{|z| \geq 1} \frac{|\phi(z)|}{|z|^2} \, \mathrm{d}z \leq \nu_{1,0}(\phi) \int_{|z| \geq 1} \frac{1}{|z|^3} \, \mathrm{d}t = 2\pi \nu_{1,0}(\phi),$$

we have now shown that $\operatorname{PV} 1/z^2 \in \mathcal{S}'(\mathbf{C})$.

For the next assertion, we let $\phi \in C_c^\infty(\mathbf{C})$. We note that by the product rule for ∂_z, we have $\partial_z(\phi/z) = (\partial_z\phi)/z - \phi/z^2$ outside of the origin. Hence, we have

$$\left\langle \partial_z \frac{1}{z}, \phi \right\rangle = -\int_{\mathbf{C}} \frac{\partial_z \phi(z)}{z}\, \mathrm{d}z = \left\langle -\operatorname{PV} \frac{1}{z^2}, \phi \right\rangle - \lim_{\varepsilon \downarrow 0} \int_{|z| \geq \varepsilon} \partial_z \left(\frac{\phi}{z}\right)(z)\, \mathrm{d}z. \tag{2.60}$$

Let $\varepsilon \in \mathbf{R}_+$ and define $\gamma : [0, 2\pi] \to \mathbf{C}$ by $\gamma(t) := \varepsilon e^{it}$. Since ϕ has compact support, it follows from Green's Integral Theorem, see Theorem C.7, that

$$\begin{aligned}\int_{|z| \geq \varepsilon} \partial_z \left(\frac{\phi}{z}\right)(z)\, \mathrm{d}z &= \frac{1}{2i} \int_\gamma \frac{\phi(z)}{z}\, \mathrm{d}\overline{z} \\ &= \frac{1}{2i} \int_0^{2\pi} \frac{\phi(\varepsilon e^{it})}{\varepsilon e^{it}} (-i\varepsilon e^{-it})\, \mathrm{d}t = -\frac{1}{2} \int_0^{2\pi} \phi(\varepsilon e^{it}) e^{-2it}\, \mathrm{d}t.\end{aligned}$$

Hence,

$$\lim_{\varepsilon \downarrow 0} \int_{|z| \geq \varepsilon} \partial_z \left(\frac{\phi}{z}\right)(z)\, \mathrm{d}z = -\frac{\phi(0)}{2} \int_0^{2\pi} e^{-2it}\, \mathrm{d}t = 0.$$

Thus, it follows from (2.60) that

$$\left\langle \partial_z \frac{1}{z}, \phi \right\rangle = \left\langle -\operatorname{PV} \frac{1}{z^2}, \phi \right\rangle$$

for all $\phi \in C_c^\infty(\mathbf{C})$. This proves the result. □

Our next order of business is to establish that the Beurling-Ahlfors transform as defined here, coincides with the definition we gave in Definition 2.42. For this, we will show that they both define the same Fourier multiplier.

2.56 Lemma. *We have* $\mathscr{F}E = -iE$.

Proof. The proof is similar to the proof of Lemma 2.40 which we refer to for more details on the arguments presented here. We note that

$$1 = \mathscr{F}\delta = \mathscr{F}(\partial_{\overline{z}} E) = \pi i \zeta \mathscr{F}E.$$

From this we conclude that the tempered distributions $\mathscr{F}E$ and $-iE$ coincide on $\mathbf{C}\backslash\{0\}$. Set $u := \mathscr{F}E + iE \in \mathcal{S}'(\mathbf{C})$. Then u is supported in the origin, which implies that u is a linear combination of derivatives of δ. Since u is homogeneous of degree -1 while $\partial^\alpha \delta$ is homogeneous of degree $-2 - |\alpha|$ for every multi-index α, we find that we must have $u = 0$. The result follows. □

This lemma implies that for all $\phi \in \mathcal{S}(\mathbf{C})$ we have

$$\mathscr{F}(\mathscr{B}\phi) = \mathscr{F}(\partial_z E)\mathscr{F}\phi = \pi i \overline{z} \mathscr{F}E \mathscr{F}\phi = \pi \overline{z} E \mathscr{F}\phi = M_{\frac{\overline{z}}{z}} \mathscr{F}\phi.$$

This means that $\mathscr{B}_2$ is an extension of $\mathscr{B}$ to $L^2(\mathbf{C})$. We conclude that $\mathscr{B}_p$ is an extension of $\mathscr{B}$ to $L^p(\mathbf{C})$ for all $p \in]1, \infty[$, as desired. From now on, we will simply denote the extensions $\mathscr{B}_p$ by $\mathscr{B}$.

We conclude this subsection with a result we will need for the Cauchy transform. For $p \in [1, \infty]$, we denote by $L_c^p(\mathbf{C})$ the classes in $L^p(\mathbf{C})$ which have a representative that has compact support. We note that the convolution $\mathscr{C}f := E * f$ makes sense for any $f \in L_c^p(\mathbf{C})$ for $p \in [1, \infty]$.

2.57 Proposition. *Let $p \in]2,\infty[$. Then $\mathscr{C}$ defines a map from $L^p_c(\mathbf{C})$ to $W^{1,p}(\mathbf{C})$ that is continuous in the sense that whenever a sequence $(f_j)_{j\in\mathbf{N}}$ in $L^p_c(\mathbf{C})$ satisfies:*

(i) $(f_j)_{j\in\mathbf{N}}$ converges in $L^p(\mathbf{C})$ to a function $f \in L^p(\mathbf{C})$,

(ii) there is a compact set $K \subseteq \mathbf{C}$ so that $\operatorname{supp} f_j \subseteq K$ for all $j \in \mathbf{N}$,

then $f \in L^p_c(\mathbf{C})$ and $(\mathscr{C} f_j)_{j\in\mathbf{N}}$ converges to $\mathscr{C} f$ in $W^{1,p}(\mathbf{C})$.

Furthermore, we have

$$\partial_{\overline{z}}\mathscr{C} f = f, \quad \partial_z \mathscr{C} f = \mathscr{B} f$$

for all $f \in L^p_c(\mathbf{C})$.

Finally, for each $f \in L^p_c(\mathbf{C})$ there exist $c, R \in \mathbf{R}_+$ so that

$$|\mathscr{C} f(z)| \leq \frac{c}{|z|} \quad \text{for a.e. } z \in \mathbf{C} \text{ satisfying } |z| \geq R.$$

Proof. For each $r \in \mathbf{R}_+$ we denote by χ_r the indicator function of the closed disc of radius r centered at the origin.

Let $f \in L^p_c(\mathbf{C})$ and pick $R \in \mathbf{R}_+$ so that the support of f is contained in the disc of radius R around the origin. First we will show that $\mathscr{C} f \in L^p(\mathbf{C})$. Let $w \in \mathbf{C}$ satisfy $|w| \leq 2R$. Then, if $|z - w| \geq 3R$ for some $z \in \mathbf{C}$, then $3R \leq |z - w| \leq |z| + 2R$, i.e., $|z| \geq R$, so that f vanishes almost everywhere for z satisfying $|z - w| \geq 3R$. This implies that

$$\mathscr{C} f(w) = \frac{1}{\pi}\int_{\mathbf{C}} \chi_{3R}(w - z)\frac{f(z)}{w - z}\,\mathrm{d}z = (\chi_{3R}E * f)(w).$$

Hence, Minkowski's inequality for convolutions, see Corollary A.5, implies that

$$\|\chi_{2R}\mathscr{C} f\|_p \leq \|\chi_{3R}E\|_1\|f\|_p. \tag{2.61}$$

Next, suppose $w \in \mathbf{C}$ satisfies $|w| > 2R$. Then, whenever $|z| \leq R$, we have $2|z| \leq 2R < |w|$ so that $|w - z| \geq |w| - |z| \geq |w|/2$. Hence,

$$|\mathscr{C} f(w)| \leq \frac{1}{\pi}\int_{|z|\leq R}\frac{|f(z)|}{|w - z|}\,\mathrm{d}z \leq \frac{2}{\pi|w|}\|f\|_1 \leq \frac{2(\pi R^2)^{\frac{1}{p'}}}{\pi|w|}\|f\|_p,$$

where in the last inequality we used Hölder's inequality. Hence,

$$\begin{aligned}\|(1 - \chi_{2R})\mathscr{C} f\|_p &\leq \frac{2(\pi R^2)^{\frac{1}{p'}}}{\pi}\|f\|_p\int_{|w|>2R}\frac{1}{|z|^p}\,\mathrm{d}z = 4(\pi R^2)^{\frac{1}{p'}}\|f\|_p\int_{2R}^{\infty} r^{1-p}\,\mathrm{d}r \\ &= \frac{4(\pi R^2)^{\frac{1}{p'}}(2R)^{2-p}}{p - 2}\|f\|_p,\end{aligned} \tag{2.62}$$

where we used $p > 2$. Thus, from (2.61) and (2.62) we conclude that

$$\mathscr{C} f = \chi_{2R}\mathscr{C} f + (1 - \chi_{2R})\mathscr{C} f \in L^p(\mathbf{C}),$$

and

$$\|\mathscr{C}f\|_p \leq k(p,R)\|f\|_p,$$

where $k(p,R) \in \mathbf{R}_+$ depends only on p and on R.

Note that Fubini's Theorem implies that $\langle \mathscr{C}f, \phi\rangle = -\langle f, \mathscr{C}\phi\rangle$ for all $\phi \in C_c^\infty(\mathbf{C})$. By (2.58) and Proposition 2.48, this implies that

$$\begin{aligned}\langle \partial_z(\mathscr{C}f), \phi\rangle &= \langle f, \mathscr{C}(\partial_z\phi)\rangle = \langle f, \mathscr{B}\phi\rangle = \langle \mathscr{B}f, \phi\rangle \\ \langle \partial_{\bar z}(\mathscr{C}f), \phi\rangle &= \langle f, \mathscr{C}(\partial_{\bar z}\phi)\rangle = \langle f, \phi\rangle,\end{aligned}$$

for all $\phi \in C_c^\infty(\mathbf{C})$. Thus, we find the distributional equalities $\partial_z(\mathscr{C}f) = \mathscr{B}f \in L^p(\mathbf{C})$, $\partial_{\bar z}(\mathscr{C}f) = f \in L^p(\mathbf{C})$. In particular, this shows us that $\mathscr{C}f \in W^{1,p}(\mathbf{C})$ with

$$\|\mathscr{C}f\|_p + \|\partial_z(\mathscr{C}f)\|_p + \|\partial_{\bar z}(\mathscr{C}f)\|_p \leq (k(p,R) + 1 + \|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))})\|f\|_p. \tag{2.63}$$

If a sequence $(f_j)_{j\in\mathbf{N}}$ in $L_c^p(\mathbf{C})$ converges in $L^p(\mathbf{C})$ to a function $f \in L^p(\mathbf{C})$ and satisfies the assumption that there is a compact set $K \subseteq \mathbf{C}$ so that $\operatorname{supp} f_j \subseteq K$ for all $j \in \mathbf{N}$, then we note that there is some $R \in \mathbf{R}_+$ so that K and hence the support of each f_j lies in the disc of radius R centered at the origin. Since convergence of a sequence in $L^p(\mathbf{C})$ implies that there is an a.e. convergent subsequence with the same limit, we conclude that f must vanish a.e. outside of K, and thus $f \in L_c^p(\mathbf{C})$. Moreover, we note that the inequality (2.63) holds for f replaced by $f - f_j$ for all $j \in \mathbf{N}$, since the support of $f - f_j$ lies in the disc of radius R centered at the origin for all $j \in \mathbf{N}$. Then, letting $j \to \infty$ proves the desired convergence.

Finally, we note that for each $f \in L_c^p(\mathbf{C})$ we have

$$\limsup_{|w|\to\infty} |w\mathscr{C}f(w)| \leq \limsup_{|w|\to\infty} \frac{1}{\pi}\int_{\mathbf{C}} \frac{|w|}{|w-z|}|f(z)|\,\mathrm{d}z \leq \frac{\|f\|_1}{\pi}$$

This proves the final assertion. □

Notes and Acknowledgments

The L^p-boundedness of the Hilbert transform dates back to the article [Ri] from 1928. The proof of this result using complex contour integration follows a section from the book *Complex Proofs of Real Theorems* by Lax and Zalcman, see [LZ, Section 3.6], and allowed us to give a complex analytical exposition of the theory in contrast to our harmonic analytical exposition which requires the language of distributions.

For our harmonic analytical approach we relied extensively on Grafakos' *Classical Fourier Analysis*, see [Gr, Chapter 4]. Our exposition of Calderón and Zygmund's Method of Rotations is adapted from the general n-dimensional results presented in this book to our 2-dimensional case. In $\mathbf{R}^n$, defining the Riesz transforms $\mathscr{R}_j$ for $j \in \{1,\dots,n\}$ as convolution with the tempered distributions $\Gamma((n+1)/2)/\pi^{(n+1)/2}\,\mathrm{PV}\,x_j/|x|^{n+1}$, one can use the Method of Rotations to show that each $\mathscr{R}_j$ is L^p-bounded for $p \in]1,\infty[$ with

$$\|\mathscr{R}_j\|_{\mathcal{L}(L^p(\mathbf{R}^n))} \leq \|\mathscr{H}\|_{\mathcal{L}(L^p(\mathbf{R}))} = \cot\frac{\pi}{2p^*}, \quad j \in \{1,\dots,n\},$$

as we have shown in Theorem 2.34 for the case $n = 2$. It is actually true that this inequality of operator norms is an equality. This is a result by Iwaniec and Martin and can be found in [IM].

Concerning homogeneous distributions such as the ones used to the define the Riesz transforms, we gave a proof of the fact that $\mathscr{F}|z|^{-1} = |\zeta|^{-1}$ in Lemma 2.40 which was based on [Gr, Section 2.4.c] and the discussion in [Gru, p. 112-114]. This is a well-known and understood result. In $\mathbf{R}^n$, for a complex parameter $a \in \mathbf{C}$ with $\operatorname{Re} a > -n$ we can define a tempered distribution u_a through the locally integrable function

$$\frac{\pi^{\frac{a+n}{2}}}{\Gamma\left(\frac{a+n}{2}\right)}|x|^a$$

in $\mathbf{R}^n$, which is a homogeneous distribution of degree a. Using analytic continuation, see [Gr, Section 2.4.c], one can extend the analytic family of tempered distributions $(u_a)_{\operatorname{Re} a > -n}$ to all $a \in \mathbf{C}$. One can then show that $\mathscr{F}u_a = u_{-n-a}$ for all $a \in \mathbf{C}$, see [Gr, Theorem 2.4.6]. Our result is then the special case $n = 2$, $a = -1$. The general result can be used to compute the Fourier transform of the Riesz transforms in $\mathbf{R}^n$, see [Gr, Exercise 4.1.10].

3 Quasiconformal Mappings and Iwaniec's Conjecture

In contrast to the previous sections, this subsection will be in a more narrative style. We will introduce the notion of quasiconformal mappings as a generalization of biholomorphisms. We will then show how this notion is related to the Beurling-Ahlfors transform. It was in the setting of quasiconformal mappings in which Iwaniec's Conjecture was first proposed by the eponymous Tadeusz Iwaniec in [Iw, Conjecture 1]. As it is not our goal to establish results in the vast theory of quasiconformal mappings, we will give references to full proofs of the results we will use. However, we will strive to emphasize results that use the Beurling-Ahlfors transform and the Cauchy transform by giving full proofs of auxiliary results where they are needed.

Geometrically speaking, holomorphic functions are interesting because they preserve angles and orientation, i.e., are conformal mappings, at points where their complex derivative doesn't vanish. For the precise definitions and results we refer to Definition C.25 and the succeeding results in Appendix C. In particular, Theorem C.31 asserts that the injective conformal mappings are precisely the biholomorphisms. A fundamental result in the study of such mappings is the Riemann Mapping Theorem.

3.1 Theorem (Riemann Mapping Theorem)**.** *Let $U \subseteq \mathbf{C}$ be non-empty, open, simply connected, and not equal to all of $\mathbf{C}$. Then, for each $z_0 \in U$ there is a biholomorphism f from U to the open unit disk satisfying $f(z_0) = 0$. Any such biholomorphism is uniquely determined up to multiplication by $e^{i\alpha}$ for $\alpha \in \mathbf{R}$.*

For now we let $U \subseteq \mathbf{C}$ be a non-empty open set. We wish to generalize the notion of conformal mappings on U to a more general setting. Rather than working in the space $C^1(U)$ of classically differentiable functions, we can choose to work in a distributional setting which leads to the use of Sobolev spaces. This immediately brings some subtleties. There exist conformal, and thus holomorphic, see Lemma C.32, mappings from $U = \mathbf{C}$ that are injective, an example being the identity map in $\mathbf{C}$. However, by Lemma 2.51 one cannot hope for such a result if one also imposes integrability of the maps. For a more fruitful theory, it is therefore sensible to consider the local Sobolev space $W^{1,1}_{loc}(U)$, consisting of those locally integrably functions whose distributional partial derivatives are also locally integrable functions.

Another subtlety is the fact that if $f \in W^{1,1}_{loc}(U)$ is an injective map satisfying $\partial_{\overline{z}} f = 0$, then the ellipticity of the linear partial differential operator $\partial_{\overline{z}}$ with constant coefficients implies that $f \in C^\infty(U)$, and is thus conformal in the classical sense, see Theorem B.37. If one wishes to generalize the notion of conformal mappings to $W^{1,1}_{loc}(U)$, this means that it may be prudent to look for a definition with a geometric flavor rather than just an analytic one.

For $f \in C^1(U)$ we write $u := \operatorname{Re} f$ and $v := \operatorname{Im} f$. Then the total derivative $Df(z)$ at a point $z \in U$, seen as a real linear map $\mathbf{R}^2 \to \mathbf{R}^2$, is given by the Jacobian matrix

$$\begin{pmatrix} \partial_x u(z) & \partial_y u(z) \\ \partial_x v(z) & \partial_y v(z) \end{pmatrix}.$$

Under the usual identification $\mathbf{R}^2 \cong \mathbf{C}$, this means that for $z \in \mathbf{C}$, $h_1, h_2 \in \mathbf{R}$ and $h = h_1 + ih_2$, we have

$$\begin{aligned} Df(z)h &= \partial_x f(z)h_1 + \partial_y f(z)h_2 = (\partial_z + \partial_{\overline{z}})f(z)h_1 + i(\partial_z - \partial_{\overline{z}})f(z)h_2 \\ &= \partial_z f(z)h + \partial_{\overline{z}} f(z)\overline{h}. \end{aligned}$$

We also recall the definition of the directional derivatives $\partial_h f$ at $z \in U$ for $h \in \mathbf{C}$ as

$$\partial_h f(z) := \lim_{t \to 0} \frac{f(z+th) - f(z)}{t} = Df(z)h.$$

The Jacobian determinant $J_f(z)$ at $z \in U$ satisfies

$$\begin{aligned} J_f(z) &= \partial_x u(z)\partial_y v(z) - \partial_y u(z)\partial_x v(z) \\ &= \frac{1}{4}((\partial_x u(z) + \partial_y v(z))^2 + (\partial_y u(z) - \partial_x v(z))^2 - (\partial_x u(z) - \partial_y v(z))^2 - (\partial_y u(z) + \partial_x v(z))^2) \\ &= |\partial_z f(z)|^2 - |\partial_{\overline{z}} f(z)|^2. \end{aligned}$$

Since the determinant of a matrix gives the oriented volume of the image under the matrix of the unit square, we say that f is orientation preserving if $J_f(z) \geq 0$ for all $z \in U$. We have just shown that this condition is equivalent to $|\partial_{\overline{z}} f(z)| \leq |\partial_z f(z)|$ for all $z \in U$. This makes sense in terms of the total derivative, since this is decomposed into the sum of multiplication by $\partial_z f$ and multiplication after conjugation by $\partial_{\overline{z}} f$. In particular, if f is holomorphic, then $Df(z)$ is simply given by multiplication with $\partial_z f = f'$.

We give another characterization of an injective conformal mapping.

3.2 Proposition. *Let $U \subseteq \mathbf{C}$ be open and let $f \in C^1(U)$ be an injective map. Then f is conformal if and only if f is orientation preserving and for all $z \in U$ the value of $|\partial_h f(z)|$ is constant for $h \in \mathbf{C}$ with $|h| = 1$.*

Proof. If f is conformal, then f is holomorphic. This means that $\partial_{\overline{z}} f(z) = 0$ for all $z \in U$. Hence, $J_f(z) = |\partial_z f(z)|^2 \geq 0$ for all $z \in U$ so that f is orientation preserving. Moreover, we find that for all $z \in U$ and all $h \in \mathbf{C}$ with $|h| = 1$ we have

$$|\partial_h f(z)| = |Df(z)h| = |\partial_z f(z)h| = |\partial_z f(z)|,$$

which is independent of h. It remains to prove the converse implication.

Let $z \in U$. If the value of $|\partial_h f(z)|$ is independent of $h \in \mathbf{C}$ with $|h| = 1$, then, by setting $h = e^{-i\alpha/2}$ for $\alpha \in \mathbf{R}$, we find that the value of

$$|\partial_h f(z)| = |Df(z)h| = |\partial_z f(z)e^{-i\frac{\alpha}{2}} + \partial_{\overline{z}} f(z)e^{i\frac{\alpha}{2}}| = |\partial_z f(z) + \partial_{\overline{z}} f(z)e^{i\alpha}|$$

is independent of $\alpha \in \mathbf{R}$. The equation

$$\partial_z f(z) + \partial_{\overline{z}} f(z)e^{i\alpha}$$

describes a circle of center $\partial_z f(z)$ and radius $|\partial_{\overline{z}} f(z)|$ as α runs through $[0, 2\pi]$. We conclude that we must either have $\partial_z f(z) = 0$ or $\partial_{\overline{z}} f(z) = 0$ for the modulus to remain constant. In the first case, we note that since f is orientation preserving, we have $|\partial_{\overline{z}} f(z)| \leq |\partial_z f(z)| = 0$ so that also $\partial_{\overline{z}} f(z) = 0$. We conclude that in either case we must have $\partial_{\overline{z}} f(z) = 0$. This implies that f is holomorphic. But an injective holomorphic function is an injective conformal mapping by Theorem C.31. The assertion follows. □

One might say that near each point $z \in U$, an injective conformal mapping on U sends circles around z to circles around z. It is this notion that we wish to generalize by allowing a quasiconformal mapping on U to send circles around a point $z \in U$ to ellipses around z that are, in a quantifiable way, not too far away from circles. Before we give a precise definition, we need to make some preliminary definitions.

For any $f \in W^{1,1}_{loc}(U)$ the distributional derivatives $\partial_z f$ and $\partial_{\overline{z}} f$ exist as functions in $L^1_{loc}(U)$. Then, for a.e. $z \in U$, this allows us to define the total derivative $Df(z) : \mathbf{C} \to \mathbf{C}$ by

$$Df(z)h := \partial_z f(z)h + \partial_{\overline{z}} f(z)\overline{h}.$$

As usual, we then write

$$\|Df(z)\|_{\mathcal{L}(\mathbf{R}^2)} = \sup_{|h|=1} |Df(z)h|.$$

For $h \in \mathbf{C}$ we may define the directional derivative $\partial_h f$ by

$$\partial_h f(z) := Df(z)h$$

for a.e. $z \in U$. Moreover, we define the Jacobian determinant J_f by

$$J_f(z) := |\partial_z f(z)|^2 - |\partial_{\overline{z}} f(z)|^2$$

for a.e. $z \in U$. Since we want J_f to be a locally integrable function, we impose the addition condition that $f \in W^{1,2}_{loc}(U)$. We say that such a function f is *orientation preserving* if $J_f(z) \geq 0$ for a.e. $z \in U$.

3.3 Definition. Let $U \subseteq \mathbf{C}$ be open and let $K \in [1, \infty[$. We say that $f \in W^{1,2}_{loc}(U)$ is K-*quasiconformal* if it is a homeomorphism onto its image, if it is orientation preserving, and if we have the inequality

$$\max_{|h|=1} |\partial_h f(z)| \leq K \min_{|h|=1} |\partial_h f(z)| \quad \text{for a.e. } z \in U. \tag{3.1}$$

Moreover, we call

$$K(f) := \inf\{K \in [1, \infty[\mid (3.1) \text{ holds}\}$$

the *maximal dilation* of f. ♢

The inequality (3.1) means that the largest distance to a point of the image of a circle around this point can, at most, be K times the shortest distance to this point. Brouwer's Invariance of Domains Theorem implies that the condition that f is a homeomorphism onto its image is equivalent to the condition that f is continuous and injective. Additionally, this theorem implies that the image of such a mapping must be open in $\mathbf{C}$. We also note that this definition of a K-quasiconformal mapping makes sense in dimensions higher than 2.

3.4 Remark. In the definition of a K-quasiconformal mapping, the condition that $f \in W^{1,2}_{loc}(U)$ is superfluous in the sense that if $f \in W^{1,1}_{loc}(U)$, then one still has that J_f is locally integrable. As a matter of fact, it is shown in [AIM, Corollary 3.3.6] that if $f \in W^{1,1}_{loc}(U)$ is orientation preserving and a homeomorphism onto its image, then

$$\int_E J_f(z)\,\mathrm{d}z \leq |f(E)| \quad \text{for all Borel measurable } E \subseteq U, \tag{3.2}$$

where $|f(E)|$ denotes the Lebesgue measure of the Borel measurable set $f(E)$. One cannot generally ask for such a result for general Lebesgue measurable sets, since the homeomorphic image of a Lebesgue measurable set need not be Lebesgue measurable.

We will show in the proof of Proposition 3.6 below that for an orientation preserving map the inequality (3.1) is equivalent to the inequality

$$\|Df(z)\|^2_{\mathcal{L}(\mathbf{R}^2)} \leq KJ_f(z) \quad \text{for a.e. } z \in U, \tag{3.3}$$

and that

$$\|Df(z)\|_{\mathcal{L}(\mathbf{R}^2)} = |\partial_z f(z)| + |\partial_{\overline{z}} f(z)|$$

for a.e. $z \in U$. Hence, by (3.2), whenever $f \in W^{1,1}_{loc}(U)$ satisfies (3.3) we must have $f \in W^{1,2}_{loc}(U)$ and f is a K-quasiconformal mapping. ♢

We call a function $f \in W^{1,2}_{loc}(U)$ quasiconformal, if there is a $K \in [1, \infty[$ so that f is K-quasiconformal. The following proposition shows us that we have indeed generalized the notion of injective conformal mappings.

3.5 Proposition. *Let $U \subseteq \mathbf{C}$ be open. Then $f \in W^{1,2}_{loc}(U)$ is 1-quasiconformal if and only if $f \in C^1(U)$ and f is an injective conformal mapping.*

Proof. If $f \in C^1(U)$ and f is an injective conformal mapping, then it follows from Proposition 3.2 that f is orientation preserving and for all $z \in U$ the value of $|\partial_h f(z)|$ is independent of $h \in \mathbf{C}$ with $|h| = 1$, which is equivalent to saying that

$$\max_{|h|=1} |\partial_h f(z)| \leq \min_{|h|=1} |\partial_h f(z)|$$

for all $z \in U$. Since injective conformal mappings are biholomorphisms, f is a homeomorphisms onto its image. Noting that $C^1(U) \subseteq W^{1,2}_{loc}(U)$, we conclude that f is 1-quasiconformal.

For the converse, we note that the proof of necessity in Proposition 3.2 also works in our more general setting to show that $\partial_{\overline{z}} f(z) = 0$ for a.e. $z \in U$. But this means that $\partial_{\overline{z}} f = 0$ in the distributional sense, which, by elliptic regularity, implies that $f \in C^\infty(U)$ and f is holomorphic, see Theorem B.37. As f is also injective, we may conclude that f is an injective conformal mapping. The assertion follows. □

It is an exercise in linear algebra to obtain different characterizations of quasiconformal mappings.

3.6 Proposition. *Let $U \subseteq \mathbf{C}$ be open, let $K \in [1, \infty[$, $\beta \in [0, 1[$ satisfy the relations*

$$K = \frac{1+\beta}{1-\beta}, \quad \beta = \frac{K-1}{K+1},$$

and let $f \in W^{1,2}_{loc}(U)$ be a homeomorphism onto its image. Then the following are equivalent:

(i) f is K-quasiconformal;

(ii) $\|Df(z)\|^2_{\mathcal{L}(\mathbf{R}^2)} \leq KJ_f(z)$ for a.e. $z \in U$;

(iii) $|\partial_{\overline{z}} f(z)| \leq \beta |\partial_z f(z)|$ *for a.e.* $z \in U$*;*

(iv) $\partial_{\overline{z}} f(z) = \mu(z)\partial_z f(z)$ *for a.e.* $z \in U$ *for some* $\mu \in L^\infty(U)$ *satisfying* $\|\mu\|_\infty \leq \beta < 1$*.*

The proof uses a lemma.

3.7 Lemma. *Let* $a, b, c, d \in \mathbf{R}$ *and let* $A : \mathbf{R}^2 \to \mathbf{R}^2$ *be a real linear map given by the matrix*

$$\begin{pmatrix} a & b \\ c & d. \end{pmatrix}$$

Then

$$\max_{|h|=1} |Ah| = \frac{1}{2}\left(\sqrt{(a+d)^2 + (c-b)^2)} + \sqrt{(a-d)^2 + (c+b)^2}\right) = |z| + |w|$$
$$\min_{|h|=1} |Ah| = \frac{1}{2}\left|\sqrt{(a+d)^2 + (c-b)^2} - \sqrt{(a-d)^2 + (c+b)^2}\right| = ||z| - |w||,$$

where

$$z = \frac{a+d}{2} + i\frac{c-b}{2}, \quad w = \frac{a-d}{2} + i\frac{c+b}{2}$$

In particular, we have

$$|\det A| = |ad - bc| = ||z|^2 - |w|^2| = \min_{|h|=1} |Ah| \max_{|h|=1} |Ah|.$$

Proof. In complex notation, for $h \in \mathbf{C}$ we have

$$Ah = zh + w\overline{h}.$$

We now have to show that $\max_{|h|=1} |Ah| = |z| + |w|$ and $\min_{|h|=1} |Ah| = ||z| - |w||$.

By the triangle inequality we have

$$||z| - |w|| \leq |Ah| \leq |z| + |w|,$$

whenever $|h| = 1$, which implies

$$\max_{|h|=1} |Ah| \leq |z| + |w|, \quad ||z| - |w|| \leq \min_{|h|=1} |Ah|.$$

It now suffices to find vectors where the maximum and minimum are attained.

If $z = |z|e^{it}$ and $w = |w|e^{is}$ for $t, s \in \mathbf{R}$, then we set $h := e^{i(s-t)/2}$. Then

$$|Ah| = ||z|e^{i\frac{s+t}{2}} + |w|e^{i\frac{s+t}{2}}| = |z| + |w|, \quad |Aih| = |izh - iw\overline{h}| = ||z| - |w||.$$

The assertion follows. □

We note that for $f \in W^{1,1}_{loc}(U)$ this lemma implies that

$$\|Df(z)\|_{\mathcal{L}(\mathbf{R}^2)} = |\partial_z f(z)| + |\partial_{\overline{z}} f(z)|$$

for a.e. $z \in U$.

Proof of Proposition 3.6. For (i)⇒(ii) we note that, by the determinant formula in Lemma 3.7, we have

$$\|Df(z)\|^2_{\mathcal{L}(\mathbf{R}^2)} = \left(\max_{|h|=1} |Df(z)h|\right)^2 \le K \min_{|h|=1} |Df(z)h| \max_{|h|=1} |Df(z)h| = KJ_f(z),$$

for a.e. $z \in U$, where we used $|J_f| = J_f$, since f is orientation preserving.

For the converse implication (ii)⇒(i), we note that the inequality in (ii) implies that f must be orientation preserving, while the inequality (3.1) is again clear by the determinant formula in Lemma 3.7.

For (ii)⇒(iii), we note that, by Lemma 3.7, for a.e. $z \in U$,

$$(|\partial_z f(z)| + |\partial_{\overline{z}} f(z)|)^2 = \|Df(z)\|^2_{\mathcal{L}(\mathbf{R}^2)} \le KJ_f(z) = K(|\partial_z f(z)|^2 - |\partial_{\overline{z}} f(z)|^2).$$

But this means that

$$|\partial_z f(z)| + |\partial_{\overline{z}} f(z)| \le K|\partial_z f(z)| - K|\partial_{\overline{z}} f(z)|$$

for a.e. $z \in U$, which is aquivalent to (iii), as desired. The implication (iii)⇒(ii) follows from a similar use of Lemma 3.7.

For (iii)⇒(iv), we define

$$\mu(z) := \begin{cases} \dfrac{\partial_{\overline{z}} f(z)}{\partial_z f(z)} & \text{if } \partial_z f(z) \ne 0 \\ 0 & \text{if } \partial_z f(z) = 0. \end{cases}$$

This satisfies the desired conditions. The implication (iv)⇒(iii) is clear. This proves the assertion. □

With f satisfying the equivalent properties in the proposition, we remark that

$$K(f) = \frac{1 + \|\mu\|_\infty}{1 - \|\mu\|_\infty}, \quad \|\mu\|_\infty = \frac{K(f) - 1}{K(f) + 1},$$

with μ as in (iv).

The equation $\partial_{\overline{z}} f = M_\mu \partial_z f$ is called the *Beltrami equation* with *Beltrami coefficient* μ. Since the Beurling-Ahlfors transform is defined on globally defined L^p functions, we will restrict our considerations to globally defined quasiconformal mappings with compactly supported Beltrami coefficients.

Let $\beta \in [0, 1[$ and let $\mu \in L^\infty_c(\mathbf{C})$ satisfy $\|\mu\|_\infty \le \beta$. The key to solving the Beltrami equation with Beltrami coefficient μ, and thus to finding quasiconformal mappings, is to invert the operator $I - M_\mu \mathscr{B}$, where I denotes the identity operator. We observe that

$$\|M_\mu \mathscr{B}\|_{\mathcal{L}(L^2(\mathbf{C}))} \le \|\mu\|_\infty \|\mathscr{B}\|_{\mathcal{L}(L^2(\mathbf{C}))} = \|\mu\|_\infty \le \beta < 1,$$

which implies that $I - M_\mu \mathscr{B}$ is invertible in $L^2(\mathbf{C})$, whose inverse is given by the Neumann series

$$(I - M_\mu \mathscr{B})^{-1} = \sum_{j \in \mathbf{Z}_{\ge 0}} (M_\mu \mathscr{B})^j,$$

where the sum converges in $\mathcal{L}(L^2(\mathbf{C}))$. As a matter of fact, we can get a better integrability exponent than $p = 2$. We will show this in Lemma 3.10 below.

This is all we will need to establish existence of solutions to the Beltrami equations under consideration.

3.8 Theorem (Existence of principal solutions to the Beltrami equation)**.** *Let $\mu \in L^\infty_c(\mathbf{C})$ with $\|\mu\|_\infty \in [0,1[$. Then there exists a unique $f \in W^{1,2}_{loc}(\mathbf{C})$ satisfying*

$$\partial_{\overline{z}} f(z) = \mu(z)\partial_z f(z) \quad \textit{for a.e. } z \in \mathbf{C},$$

with the condition that there exist $c, R \in \mathbf{R}_+$ so that

$$|f(z) - z| \leq \frac{c}{|z|} \quad \textit{for a.e. } z \in \mathbf{C} \textit{ satisfying } |z| \geq R. \tag{3.4}$$

We call the solution $f \in W^{1,2}_{loc}(\mathbf{C})$ as in the theorem the *principal solution* to the Beltrami equation with Beltrami coefficient $\mu \in L^\infty_c(\mathbf{C})$.

Proof. We will first find a solution $g \in W^{1,2}(\mathbf{C})$ to the equation

$$\partial_{\overline{z}} g(z) = \mu(z)\partial_z g(z) + \mu(z) \quad \text{for a.e. } z \in \mathbf{C}. \tag{3.5}$$

This equation can be rewritten as

$$(I - M_\mu \mathscr{B})(\partial_{\overline{z}} g) = \mu.$$

As $\mu \in L^\infty_c(\mathbf{C}) \subseteq L^2(\mathbf{C})$, we can define

$$\omega := (I - M_\mu \mathscr{B})^{-1}\mu = \sum_{j \in \mathbf{Z}_{\geq 0}} (M_\mu \mathscr{B})^j \mu = \mu + \mu \sum_{j \in \mathbf{N}} \mathscr{B}(M_\mu \mathscr{B})^{j-1}\mu \in L^2_c(\mathbf{C}).$$

By Proposition 2.57 we can set $g := \mathscr{C}\omega \in W^{1,2}(\mathbf{C})$. By retracing our steps and by noting that $\partial_{\overline{z}} g = \omega$, we note that g is indeed a solution to (3.5). Moreover, by Proposition 2.57 we can find $c, R \in \mathbf{R}_+$ so that

$$|g(z)| \leq \frac{c}{|z|} \quad \text{for a.e. } z \in \mathbf{C} \text{ satisfying } |z| \geq R.$$

Now set $f := z + g \in W^{1,2}_{loc}(\mathbf{C})$. Then f satisfies (3.4). Moreover, we have

$$\partial_{\overline{z}} f(z) = \partial_{\overline{z}} g(z) = \mu(z)\partial_z g(z) + \mu(z)\partial_z z = \mu(z)\partial_z f(z) \quad \text{for a.e. } z \in \mathbf{C}.$$

This establishes existence.

For uniqueness, suppose $\tilde{f} \in W^{1,2}_{loc}(\mathbf{C})$ is another solution. Then $\tilde{g} := \tilde{f} - z \in W^{1,2}_{loc}(\mathbf{C})$ is a solution to the equation (3.5). Set $h := \tilde{f} - f = \tilde{g} - g$. Since μ is compactly supported and by the decay properties of $\tilde{f}$ and f, we can find $c', R' \in \mathbf{R}_+$ so that

$$|h(z)| \leq \frac{c'}{|z|} \text{ and } \partial_{\overline{z}} h(z) = 0 \quad \text{for a.e. } z \in \mathbf{C} \text{ satisfying } |z| \geq R'. \tag{3.6}$$

Since $\partial_{\bar{z}} h \in L^2_{loc}(\mathbf{C})$ vanishes outside of a compact set, we conclude that $\partial_{\bar{z}} h \in L^2(\mathbf{C})$. But then

$$(I - M_\mu \mathscr{B})(\partial_{\bar{z}} h) = 0$$

so that $\partial_{\bar{z}} h = 0$ in $L^2(\mathbf{C})$ by invertibility of $I - M_\mu \mathscr{B}$ in $L^2(\mathbf{C})$. Since (3.6) implies that h is bounded, we conclude from Liouville's Theorem, Lemma 2.51, that h is constant. But then the decay property (3.6) implies that $h = 0$, as desired. □

The principal solution we have found above actually turns out to be a homeomorphism of the plane and, in particular, is a quasiconformal mapping of maximal dilation $(1 + \|\mu\|_\infty)/(1 - \|\mu\|_\infty)$. A full proof of this result can be found in [AIM, Theorem 5.3.2]. We also wish to state a result that that is referred to as the Measurable Riemann Mapping Theorem and was established by Lars Ahlfors and Lipman Bers in 1960 in [AB]. We state it here without proof. Here we denote the Riemann sphere by $\hat{\mathbf{C}} := \mathbf{C} \cup \{\infty\}$.

3.9 Theorem (Measurable Riemann Mapping Theorem)**.** *Let $\mu \in L^\infty_c(\mathbf{C})$ with $\|\mu\|_\infty \in [0, 1[$. Then there is a unique homeomorphism $f : \hat{\mathbf{C}} \to \hat{\mathbf{C}}$ satisfying*

$$\partial_{\bar{z}} f(z) = \mu(z) \partial_z f(z) \quad \text{for a.e. } z \in \mathbf{C}$$

that is a quasiconformal mapping that fixes the points 0, 1, and ∞.

We note in particular that this theorem establishes existence of quasiconformal mappings of any maximal dilation in domains $U \subseteq \mathbf{C}$. Indeed, by extending the Beltrami coefficient of the corresponding Beltrami equation by 0 outside of U, one obtains a Beltrami coefficient as in the theorem. Restricting the solution from the theorem to U yields the desired quasiconformal mapping.

Next, we will establish higher integrability results for solutions of Beltrami equations. This was the phenomenon, as perhaps first observed by Bojarski, see [Bo], that Iwaniec was studying which led him to his conjecture. In Bojarski's work he used the recently developed interpolation techniques to obtain continuity results of integrability exponents from which higher integrability results for solutions to certain partial differential equations can be found. A prime example is the following lemma, used in conjunction with the Neumann series argument we have used so far.

3.10 Lemma. *Let $\beta \in [0, 1[$. Then there exists an $\varepsilon \in \mathbf{R}_+$ so that for all $p \in]2 - \varepsilon, 2 + \varepsilon[$ we have*

$$\beta \|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} < 1.$$

Proof. The Riesz-Thorin Interpolation Theorem implies that the function $t \mapsto \log \|\mathscr{B}\|_{\mathcal{L}(L^{t^{-1}}(\mathbf{C}))}$ is convex in $]0, 1[$. Since such functions are continuous, we conclude that the function $p \mapsto \|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} = e^{\log \|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))}}$ is continuous in $]1, \infty[$. But then, since $\|\mathscr{B}\|_{\mathcal{L}(L^2(\mathbf{C}))} = 1$, we conclude from the fact that $\beta \|\mathscr{B}\|_{\mathcal{L}(L^2(\mathbf{C}))} < 1$, that there must be some open interval around 2 where this condition still holds. This proves the assertion. □

Let $\beta \in [0, 1[$. Then we can set

$$p(\beta) := \sup\{p \in]1, \infty[\mid I - M_\mu \mathscr{B} \text{ is invertible in } L^p(\mathbf{C}) \text{ for all } \mu \in L^\infty(\mathbf{C}) \text{ satisfying } \|\mu\|_\infty \leq \beta\}.$$

An argument using Neumann series as above yields

$$\sup\{p \in]1,\infty[\mid \beta\|\mathscr{B}\||_{\mathcal{L}(L^p(\mathbf{C}))} < 1\} \leq p(\beta), \tag{3.7}$$

which means that by the lemma we have $p(\beta) \in]2,\infty]$.

Let $U \subseteq \mathbf{C}$ be open. Bojarski showed that any solution $f \in W^{1,2}_{loc}(U)$ to the Beltrami equation with Beltrami coefficient $\mu \in L^\infty(U)$ with $\|\mu\|_\infty \leq \beta < 1$ must actually lie in $W^{1,p}_{loc}(U)$ for some $p \in]2,\infty[$. This is a consequence of the following result, which we call Bojarski's Theorem.

3.11 Theorem (Bojarski's Theorem)**.** *Let $U \subseteq \mathbf{C}$ be open, let $\beta \in [0,1[$, and let $\mu \in L^\infty(U)$ satisfy $\|\mu\|_\infty \leq \beta$. If $f \in W^{1,2}_{loc}(U)$ satisfies*

$$\partial_{\overline{z}} f(z) = \mu(z)\partial_z f(z) \quad \textit{for a.e. } z \in U,$$

then $f \in W^{1,p}_{loc}(U)$ for all $p \in [2, p(\beta)[$.

For the proof we will use the following version of the Sobolev Embedding Theorem:

3.12 Theorem. *For all $p \in [2,\infty[$ we have the continuous inclusion $W^{1,2}(\mathbf{C}) \subseteq L^p(\mathbf{C})$.*

Proof. Let $\phi \in C^\infty_c(\mathbf{C})$ be arbitrary and let $\chi \in C^\infty_c(\mathbf{R})$ satisfy $\chi(\mathbf{R}) \subseteq [0,1]$, $\chi(0) = 1$, and $\operatorname{supp}\chi \subseteq]-1,1[$. For $\alpha \in \mathbf{R}$, the chain rule implies

$$\phi(0) = -\int_0^1 \partial_r(\chi(r)\phi(re^{i\alpha}))\,\mathrm{d}r = -\int_0^\infty \left(\chi(r)D\phi(re^{i\alpha})e^{i\alpha} + \chi'(r)\phi(re^{i\alpha})\right)\,\mathrm{d}r.$$

Hence, by employing polar coordinates,

$$\begin{aligned} -2\pi\phi(0) &= \int_0^1 r\int_0^{2\pi} \frac{1}{r}\left(\chi(r)D\phi(re^{i\alpha})e^{i\alpha} + \chi'(r)\phi(re^{i\alpha})\right)\,\mathrm{d}\alpha\,\mathrm{d}r \\ &= \int_D \frac{1}{|z|}\left(\chi(|z|)D\phi(z)\frac{z}{|z|} + \chi'(|z|)\phi(z)\right)\,\mathrm{d}z, \end{aligned}$$

where $D \subseteq \mathbf{C}$ denotes the open unit disk. Since $D\phi(z)h = \partial_z\phi(z)h + \partial_{\overline{z}}\phi(z)\overline{h}$, this implies that there is some $c \in \mathbf{R}_+$ so that

$$|\phi(0)| \leq c\int_D \frac{1}{|z|}\left(|\phi(z)| + |\partial_z\phi(z)| + |\partial_{\overline{z}}\phi(z)|\right)\,\mathrm{d}z. \tag{3.8}$$

Denote by χ_D the indicator function of D and fix $z_0 \in \mathbf{C}$. If we replace ϕ in (3.8) by $z \mapsto \phi(z_0 - z)$, then we obtain

$$|\phi(z_0)| \leq c\int_D \frac{1}{|z|}\left(|\phi(z_0-z)| + |\partial_z\phi(z_0-z)| + |\partial_{\overline{z}}\phi(z_0-z)|\right)\,\mathrm{d}z = c\left(\frac{\chi_D}{|z|} * \psi\right)(z_0), \tag{3.9}$$

where $\psi := |\phi| + |\partial_z\phi| + |\partial_{\overline{z}}\phi|$.

Let $p \in [2,\infty[$ and let $q \in [1,\infty[$ satisfy

$$\frac{1}{p} + 1 = \frac{1}{q} + \frac{1}{2}.$$

Then

$$q = 2 - \frac{4}{2+p} < 2,$$

so that $\chi_D/|z| \in L^q(\mathbf{C})$. Using (3.9) and Young's inequality for convolutions, see Theorem A.3, we obtain

$$\|\phi\|_p \leq c \left\| \frac{\chi_D}{|z|} * \psi \right\|_p \leq c \left\| \frac{\chi_D}{|z|} \right\|_q \|\psi\|_2.$$

Since $f \mapsto \||f|+|\partial_z f|+|\partial_{\overline{z}} f|\|_2$ defines a norm on $W^{1,2}(\mathbf{C})$, equivalent to the usual one, we conclude from the definition of ψ that there is some constant $c' \in \mathbf{R}_+$, independent of ϕ, so that

$$\|\phi\|_p \leq c' \|\phi\|_{W^{1,2}(\mathbf{C})}.$$

The result now follows from density of $C_c^\infty(\mathbf{C})$ in $W^{1,2}(\mathbf{C})$. □

Proof of Bojarski's Theorem. By Lemma 1.6 we need to check that for all $\phi \in C_c^\infty(\mathbf{C})$ we have $\phi f \in W^{1,p}(U)$ whenever $p \in [2, p(\beta)[$. Let $\phi \in C_c^\infty(U)$. Then $\phi f \in W^{1,2}(U)$ extends to a function $g \in W^{1,2}(\mathbf{C})$ by declaring that it vanishes outside of the support of ϕ. We also extend μ to $\mathbf{C}$ by 0. Then, setting $\psi := (\partial_{\overline{z}}\phi - \mu\partial_z\phi)f$, we have

$$\begin{aligned}
\partial_{\overline{z}} g(z) &= \phi(z)\partial_{\overline{z}} f(z) + \partial_{\overline{z}}\phi(z) f(z) \\
&= \mu(z)\phi(z)\partial_z f(z) + \mu(z)\partial_z\phi(z) f + \psi(z) \\
&= \mu(z)\partial_z g(z) + \psi(z)
\end{aligned}$$

for a.e. $z \in \mathbf{C}$. By Theorem 2.46, this implies that

$$(I - M_\mu \mathscr{B})(\partial_{\overline{z}} g) = \psi. \tag{3.10}$$

It follows from Theorem 3.12 that $g \in L^p(\mathbf{C})$ for all $p \in [2, \infty[$. Similarly, for all $p \in [2, \infty[$ we find that $(\partial_{\overline{z}}\phi)f, (\partial_z\phi)f \in L^p(\mathbf{C})$ so that $\psi \in L^p(\mathbf{C})$. If $p \in [2, p(\beta)[$, then the invertability of $I - M_\mu\mathscr{B}$ in (3.10) in $L^p(\mathbf{C})$ implies that $\partial_{\overline{z}} g \in L^p(\mathbf{C})$. But then also $\partial_z g = \mathscr{B}(\partial_{\overline{z}} g) \in L^p(\mathbf{C})$. In conclusion, we have $g, \partial_z g, \partial_{\overline{z}} g \in L^p(\mathbf{C})$ whenever $p \in [2, p(\beta)[$. This proves that $g \in W^{1,p}(\mathbf{C})$ and thus $\phi f \in W^{1,p}(U)$ whenever $p \in [2, p(\beta)[$. The assertion follows. □

If, for $\beta \in [0, 1[$ and $U \subseteq \mathbf{C}$ open, we set $K = (1+\beta)/(1-\beta)$ and

$$P(\beta, U) := \sup\{p \in [2, \infty[\mid \text{ any } K\text{-quasiconformal mapping in } U \text{ lies in } W^{1,p}_{loc}(U)\},$$

then we note that it follows from Bojarski's Theorem and (3.7) that we have the chain of inequalities

$$\sup\{p \in]1, \infty[\mid \beta\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} < 1\} \leq p(\beta) \leq P(\beta, U). \tag{3.11}$$

By exhibiting a specific example we can determine an upper bound for $P(\beta, U)$.

3.13 Lemma. *Let $U \subseteq \mathbf{C}$ be open and $\beta \in [0, 1[$. Then $p(\beta, U) \leq 1 + 1/\beta \in]2, \infty]$.*

Proof. Set $K := (1+\beta)/(1-\beta)$ and define $\rho_1, \rho_2 : \mathbf{R}_{\geq 0} \to \mathbf{R}_{\geq 0}$ by $\rho_1(t) := t^{1/K}$ and $\rho_2(t) := t^K$. Then the corresponding radial stretchings $f_1(z) = \rho_1(|z|)z/|z| = z|z|^{1/K-1}$, $f_2(z) = \rho_2(|z|)z/|z| = z|z|^{K-1}$ are continuous in $\mathbf{C}$ and inverses of each other. By Example 2.49, we have

$$\partial_z f_1(z) = \frac{\frac{1}{K}+1}{2}|z|^{\frac{1}{K}-1}, \quad \partial_{\bar{z}} f_1(z) = \frac{z}{\bar{z}}\frac{\frac{1}{K}-1}{2}|z|^{\frac{1}{K}-1}$$

so that $\partial_z f_1, \partial_{\bar{z}} f \in L^p_{loc}(\mathbf{C})$ whenever $p(1/K-1) > -2$ and $\partial_z f_1, \partial_{\bar{z}} f \notin W^{1,p}_{loc}(\mathbf{C})$ when $p(1/K-1) \leq -2$. This means that

$$\begin{aligned} f_1 \in W^{1,p}_{loc}(\mathbf{C}) \text{ if } p < \frac{2K}{K-1} = 1 + \frac{1}{\beta} \\ f_1 \notin W^{1,p}_{loc}(\mathbf{C}) \text{ if } p \geq \frac{2K}{K-1} = 1 + \frac{1}{\beta}. \end{aligned} \tag{3.12}$$

Moreover, we note that f_1 is K-quasiconformal, since

$$|\partial_{\bar{z}} f(z)| = \left| \frac{z}{\bar{z}} \frac{\frac{1}{K}-1}{\frac{1}{K}+1} \partial_z f(z) \right| = \beta |\partial_z f(z)|$$

for a.e. $z \in \mathbf{C}$. By picking $w \in U$ and by considering the restriction of $z \mapsto f_1(z-w)$ to U, we conclude from (3.12) that $P(\beta, U) \leq 1 + 1/\beta$. $\square$

We make a particular note that Lemma 3.13 and (3.11) implies that

$$\|\mathscr{B}\|_{\mathcal{L}(L^{1+1/\beta}(\mathbf{C}))} \geq \frac{1}{\beta}$$

for $\beta \in]0,1[$, giving another proof of the upper bound $\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} \geq p^* - 1$ for $p \in]1, \infty[$. In fact, we now see that Iwaniec's Conjecture would imply that

$$\sup\{p \in]1,\infty[\mid \beta\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} < 1\} = p(\beta) = P(\beta, U) = 1 + \frac{1}{\beta}$$

for any open $U \subseteq \mathbf{C}$ and $\beta \in [0, 1[$. Remarkably, it has actually been proven that $P(\beta, U) = 1+1/\beta$. Goldstein announced a proof of this result in [Go] in 1980. It were precisely these considerations that led Iwaniec to his conjecture in 1982 in the first place. As a matter of fact, Iwaniec showed that the equality $P(\beta, U) = 1 + 1/\beta$ is an immediate consequence of the conjecture

$$\lim_{p\to\infty} \frac{\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))}}{p} = 1, \tag{3.13}$$

which is an immediate consequence of Iwaniec's Conjecture. Indeed, using existence results of quasiconformal mappings and area distortion results for such mappings he showed in [Iw, Theorem 5] that

$$P(\beta, U) \geq \left(1 + \frac{1}{\beta}\right) 2^{1-a},$$

where

$$a := \liminf_{p\to\infty} \frac{\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))}}{p}.$$

Results of this kind were shown by Gehring, and Reich in [GR] from 1966, following the work of Bojarski. They conjectured bounds of the distortion of area under quasiconformal mappings which imply the equality $P(\beta, U) = 1 + 1/\beta$. These conjectures were later shown to be true by Astala in [As] in 1994.

In the following, we write $|E|$ for the Lebesgue measure of a measurable set $E \subseteq \mathbf{C}$.

3.14 Theorem (Astala, 1994)**.** *Let $K \in [1,\infty[$ and let $D \subseteq \mathbf{C}$ denote the open unit disk. Then there is a constant $c(K) \in \mathbf{R}_+$, depending only on K, so that for all K-quasiconformal mappings $f : D \to D$ that satisfy $f(0) = 0$ we have*

$$|f(E)| \leq c(K)|E|^{\frac{1}{K}},$$

for all Borel measurable $E \subseteq D$.

3.15 Corollary. *Let $U \subseteq \mathbf{C}$ be open and let $\beta \in [0, 1[$. Then $P(\beta, U) = 1 + 1/\beta$.*

We conclude this section by sketching the proof of this corollary. We note, by Lemma 3.13, that it suffices to show that

$$P(\beta, U) \geq 1 + \frac{1}{\beta} = \frac{2K}{K-1} \in]2, \infty],$$

where $K := (1 + \beta)/(1 - \beta) \in [1,\infty[$. Picking neighborhoods of disks in U, using the Riemann Mapping Theorem on this neighborhood and the image of this neighborhood under a map, we can compose this map with biholomorphisms to see that it suffices to consider integrability of K-quasiconformal mappings of the open unit disk D to itself that fix the origin. Let f be such a mapping. Then observe that, as in Remark 3.4, we have, setting $E_t = \{z \in D \mid J_f(z) \geq t\}$ for $t \in \mathbf{R}_+$, that

$$t|E_t| \leq \int_{E_t} J_f(z)\,\mathrm{d}z \leq |f(E_t)| \leq c|E_t|^{\frac{1}{K}},$$

where $c \in \mathbf{R}_+$ is as in the theorem. This implies that

$$|E_t| \leq \left(\frac{c}{t}\right)^{\frac{K}{K-1}}. \tag{3.14}$$

We remark that this actually establishes that J_f lies in the weak $L^{K/(K-1)}$ space of unit disk.

Since f is continuous, it certainly lies in $L^p(D)$ for any $p \in [1,\infty]$. We conclude from Proposition 3.6 that

$$\|Df(z)\|^2_{\mathcal{L}(\mathbf{R}^2)} \leq KJ_f(z) \quad \text{for a.e. } z \in D.$$

Since

$$\|Df(z)\|_{\mathcal{L}(\mathbf{R}^2)} = |\partial_z f(z)| + |\partial_{\bar{z}} f(z)|$$

for a.e. $z \in D$, it now suffices to show that $J_f \in L^p(D)$ for any $p \in [2, K/(K-1)[$. For such a p we find, by (3.14), that

$$\begin{aligned}\int_D |J_f(z)|^p \,\mathrm{d}z &= p \int_{\mathbf{R}_+} t^{p-1} |E_t| \,\mathrm{d}t \\ &\leq p|D| \int_0^1 t^{p-1} \,\mathrm{d}t + c^{\frac{K}{K-1}} p \int_1^\infty t^{p-1} t^{-\frac{K}{K-1}} \,\mathrm{d}t < \infty,\end{aligned}$$

where we note that the second integral is finite since $p - K/(K-1) - 1 < -1$. This proves the desired result.

Historical Notes Regarding Iwaniec's Conjecture

First we wish to briefly discuss the conjecture (3.13). The best known result in this direction so far is the estimate

$$\limsup_{p \to \infty} \frac{\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))}}{p} \leq \sqrt{2}. \tag{3.15}$$

Using martingale techniques, based on Burkholder's work in 1984 in obtaining optimal constants for estimates for certain martingale transforms (which, remarkably, is the constant $p^* - 1$, see [Bu]), it was shown by Dragičević and Volberg in [DV] from 2005 that we have the inequality

$$\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq \sqrt{2}(p-1) \left(\frac{1}{2\pi} \int_0^{2\pi} |\cos t|^p \,\mathrm{d}t \right)^{-\frac{1}{p}} \quad \text{for } p \in [2, \infty[,$$

which implies (3.15). In 2008 it was shown by Bañuelos and Janakiraman in [BJ] that we have the (asymptotically better) estimate

$$\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq \sqrt{2p(p-1)} \quad \text{for } p \in [2, \infty[.$$

Noting that this gives the estimate $\|\mathscr{B}\|_{\mathcal{L}(L^2(\mathbf{C}))} \leq 2$ while we know that $\|\mathscr{B}\|_{\mathcal{L}(L^2(\mathbf{C}))} = 1$, they showed, using the Riesz-Thorin Interpolation Theorem, that this gives the estimate

$$\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq 1.575(p^* - 1) \quad \text{for } p \in]1, \infty[.$$

So much for our historical discussion concerning partial results regarding Iwaniec's conjectures.

4 The Burkholder Functional

4.1 Motivation: Integral Estimates of the Jacobian Determinant

In the previous section we have seen how area distortion estimates lead to the local integrability of the Jacobian determinant of quasiconformal mappings. This led to the fact that any K-quasiconformal mapping in $U \subseteq \mathbf{C}$ lies in $W^{1,p}_{loc}(U)$ for $2 \leq p < 2K/(K-1)$, which, in turn, served as a motivation for Iwaniec's Conjecture. We will now discuss general estimates involving the Jacobian determinant of a map $f \in W^{1,p}(\mathbf{C})$ for $p \in]1,\infty[$ in an attempt to prove the validity of Iwaniec's Conjecture.

We will denote the set of 2×2 matrices with real coefficients by $\mathbf{R}^{2\times 2}$. Moreover, for any $A \in \mathbf{R}^{2\times 2}$ we will denote its operator norm by

$$|A| := \|A\|_{\mathcal{L}(\mathbf{R}^2)} = \sup_{|h|=1} |Ah|,$$

where we identify a matrix with its corresponding linear operator. Then, for any $f \in W^{1,1}_{loc}(\mathbf{C})$, we have

$$|Df(z)| = |\partial_z f(z)| + |\partial_{\overline{z}} f(z)|, \quad J_f(z) = |\partial_z f(z)|^2 - |\partial_{\overline{z}} f(z)|^2 = |Df(z)|(|\partial_z f(z)| - |\partial_{\overline{z}} f(z)|)$$

for a.e. $z \in \mathbf{C}$.

Let us take a step back for the moment and attempt to establish the upper bound

$$\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq p^* - 1$$

using the same strategy we used to prove the upper bound for the Hilbert transform, mutatis mutandis. We started the proof of the upper bound for the Hilbert transform by using Pichorides' inequality from Proposition 2.27. This was done because the left-hand side in (2.43) is a subharmonic function. In a sense, subharmonicity is a form of convexity. Indeed, it is a direct generalization of the concept of convexity in one dimension. Due to the work of Burkholder in [Bu2], we have a similar inequality in our current situation.

4.1 Theorem. *Let $p \in]1,\infty[$, $a, b \in \mathbf{R}$, and suppose $L \in [1,\infty[$ satisfies $L \geq p^* - 1$. Then*

$$p\left(\frac{L}{1+L}\right)^{p-1}(L|a| - |b|)(|a| + |b|)^{p-1} \leq L^p|a|^p - |b|^p. \tag{4.1}$$

We will see that the left-hand side of (4.1) has (conjecturally nice) convexity properties, which we will attempt to utilize.

Proof. We will show that

$$p\left(\frac{L}{1+L}\right)^{p-1}(L|a| - |b|) \leq L^p|a|^p - |b|^p \quad \text{if } |a| + |b| = 1. \tag{4.2}$$

This is sufficient, as the general case follows by passing to $a/(|a|+|b|)$, $b/(|a|+|b|)$ for arbitrary $a, b \in \mathbf{R}$ that are not both 0.

We define

$$\phi : [0,1] \to \mathbf{R}, \quad \phi(x) := L^p x^p - (1-x)^p - p\left(\frac{L}{1+L}\right)^{p-1}(Lx + x - 1).$$

If we can now show that $\phi(x) \geq 0$ for all $x \in [0,1]$, then (4.2) follows by noting that $\phi(|a|) \geq 0$ for $a, b \in \mathbf{R}$, with $|a| + |b| = 1$.

For the case $p = 2$, we have

$$\phi(x) = \frac{L-1}{L+1}(Lx + x - 1)^2 \geq 0 \quad \text{for all } x \in [0,1].$$

Now assume $p \neq 2$. Taking derivatives, we find

$$\phi'(x) = pL^p x^{p-1} + p(1-x)^{p-1} - p\frac{L^{p-1}}{(1+L)^{p-2}},$$
$$\phi''(x) = p(p-1)(L^p x^{p-2} - (1-x)^{p-2})$$

for $x \in]0,1[$. Then we note that ϕ'' has a unique zero at $x = (L^{p/(p-2)} + 1)^{-1} \in]0,1[$. But, by Rolle's Theorem, this means that ϕ' can have at most two zeroes in $]0,1[$. Since these cannot both be local minima of ϕ, we conclude that ϕ can have at most one local minimum in $]0,1[$.

Since $1 - 1/(1+L) = L/(1+L)$, we find that

$$\phi\left(\frac{1}{1+L}\right) = \phi'\left(\frac{1}{1+L}\right) = 0, \quad \phi''\left(\frac{1}{1+L}\right) = p(p-1)\frac{L^{p-2}}{(1+L)^{p-2}}(L^2 - 1) > 0,$$

where the last inequality follows from the fact that $p \neq 2$ implies that $L > 1$. We conclude that $1/(1+L)$ is the point at which ϕ attains its unique local minimum in $]0,1[$, where ϕ attains the value 0. If we can now show that ϕ is non-negative at the endpoints of $[0,1]$, then we can conclude that $\phi(x) \geq 0$ for all $x \in [0,1]$, as desired.

Note that, since $t \mapsto t/(t+1) = 1 - 1/(t+1)$ is an increasing function, we have

$$\phi(0) = p\left(\frac{L}{1+L}\right)^{p-1} - 1 \geq p\left(\frac{p^*-1}{p^*}\right)^{p-1} - 1,$$
$$\phi(1) = L^p\left(1 - \frac{p}{(1+L)^{p-1}}\right) \geq L^p\left(1 - \frac{p}{(p^*)^{p-1}}\right).$$

Thus, we have to show that

$$p\left(\frac{p^*-1}{p^*}\right)^{p-1} \geq 1, \quad \frac{p}{(p^*)^{p-1}} \leq 1. \tag{4.3}$$

We consider the two cases $p \in]1,2[$ and $p \in]2,\infty[$.

First assume that $p \in]1,2[$. Then $p^* = p/(p-1)$. Hence,

$$p\left(\frac{p^*-1}{p^*}\right)^{p-1} = p^{2-p} > 1,$$

which establishes the first inequality in (4.3). For the second inequality, we recall Young's inequality

$$rs \leq \frac{r^q}{q} + \frac{s^{q'}}{q'} \quad \text{for } r, s \in \mathbf{R}_{\geq 0},\ q \in]1, \infty[. \tag{4.4}$$

Applying this to $r = p^{2-p}$, $s = (p-1)^{p-1}$, $q = 1/(2-p)$, $q' = q/(q-1) = 1/(p-1)$, we obtain

$$\frac{p}{(p^*)^{p-1}} = p^{2-p}(p-1)^{p-1} \leq (2-p)p + (p-1)(p-1) = 1,$$

as desired.

Now assume that $p \in]2, \infty[$. Then the second inequality in (4.3) is shown by noting that

$$\frac{p}{(p^*)^{p-1}} = \frac{1}{p^{p-2}} < 1.$$

For the first inequality we use (4.4) with $r = p^{(p-2)/(p-1)}$, $s = 1$, $q = (p-1)/(p-2)$, $q' = q/(q-1) = p-1$ to find

$$p^{\frac{p-2}{p-1}} \leq \frac{p-2}{p-1}p + \frac{1}{p-1} = p-1$$

so that

$$p\left(\frac{p^*-1}{p^*}\right)^{p-1} = \frac{(p-1)^{p-1}}{p^{p-2}} \geq 1.$$

The assertion follows. □

In the case of Pichorides' result, we proceeded to show that the integral over the left-hand side in (2.43) was non-negative for $a = u$ and $b = \mathscr{H}u$ for a real-valued $u \in C_c^\infty(\mathbf{R})$. Fixing $p \in]1, \infty[$ and assuming this strategy, we recall by Proposition 2.50 that it is sufficient to consider the case where $a = \partial_z \phi$ and $b = \partial_{\overline{z}} \phi$ for $\phi \in C_c^\infty(\mathbf{C})$ (or in any space between $C_c^\infty(\mathbf{C})$ and $\mathring{W}^{1,p}(\mathbf{C})$) in (4.1). In fact, we wish to show that

$$\int_{\mathbf{C}} ((p^*-1)|\partial_z \phi(z)| - |\partial_{\overline{z}} \phi(z)|)(|\partial_z \phi(z)| + |\partial_{\overline{z}} \phi(z)|)^{p-1} \, \mathrm{d}z \geq 0. \tag{4.5}$$

Then, indeed, it follows from Theorem 4.1 with $L = p^* - 1$ that

$$(p^*-1)^p \|\partial_z \phi(z)\|_p^p - \|\partial_{\overline{z}} \phi(z)\|_p^p \geq 0,$$

and thus $\|\mathscr{B}\|_{\mathcal{L}(L^p(\mathbf{C}))} \leq p^* - 1$, verifying Iwaniec's Conjecture. In an attempt to verify the inequality (4.5), we note that it can be rewritten as

$$\int_{\mathbf{C}} |D\phi(z)|^{p-2} J_\phi(z) \, \mathrm{d}z \leq \frac{p^*-2}{p^*} \int_{\mathbf{C}} |D\phi(z)|^p \, \mathrm{d}z = \left|1 - \frac{2}{p}\right| \int_{\mathbf{C}} |D\phi(z)|^p \, \mathrm{d}z.$$

We can then formulate a stronger conjecture than Iwaniec's Conjecture.

4.2 Conjecture. *Let $p \in]1,\infty[$. Then*

$$\int_{\mathbf{C}} ((p^*-1)|\partial_{\overline{z}}\phi| - |\partial_z\phi(z)|)(|\partial_z\phi(z)| + |\partial_{\overline{z}}\phi(z)|)^{p-1}\,\mathrm{d}z \geq 0,$$

or equivalently,

$$\int_{\mathbf{C}} |D\phi(z)|^{p-2}J_\phi(z)\,\mathrm{d}z \leq \left|1-\frac{2}{p}\right| \int_{\mathbf{C}} |D\phi(z)|^p\,\mathrm{d}z,$$

for all $\phi \in C_c^\infty(\mathbf{C})$.

The case $p=2$ is easily verifiable. Pick $\phi \in C_c^\infty(\mathbf{C})$. Then, by partially integrating twice, we obtain

$$\int_{\mathbf{C}} |\partial_z\phi(z)|^2\,\mathrm{d}z = \int_{\mathbf{C}} \partial_z\phi(z)\partial_{\overline{z}}\overline{\phi}(z)\,\mathrm{d}z = \int_{\mathbf{C}} \partial_{\overline{z}}\phi(z)\partial_z\overline{\phi}(z)\,\mathrm{d}z = \int_{\mathbf{C}} |\partial_{\overline{z}}\phi(z)|^2\,\mathrm{d}z,$$

or equivalently,

$$\int_{\mathbf{C}} J_\phi(z)\,\mathrm{d}z = 0.$$

By Proposition 2.50, this gives us another proof of the fact that $\|\mathscr{B}\|_{\mathcal{L}(L^2(\mathbf{C}))} = 1$. We will prove the following partial result:

4.3 Proposition. *Let $p \in]1,\infty[$. Then there exists an $L \in [1,\infty[$ with $L \geq p^*-1$ so that*

$$\int_{\mathbf{C}} (L|\partial_{\overline{z}}f| - |\partial_z f(z)|)(|\partial_z f(z)| + |\partial_{\overline{z}}f(z)|)^{p-1}\,\mathrm{d}z \geq 0,$$

or equivalently,

$$\int_{\mathbf{C}} |Df(z)|^{p-2}J_f(z)\,\mathrm{d}z \leq \frac{L-1}{L+1}\int_{\mathbf{C}} |Df(z)|^p\,\mathrm{d}z,$$

for all $f \in W^{1,p}(\mathbf{C})$. More precisely, this result holds for $L = \|\mathscr{B}\|^p_{\mathcal{L}(L^p(\mathbf{C}))}$.

We note that Conjecture 4.2 is the validity of the proposition for $L = p^*-1$. For the proof we will use the following lemma:

4.4 Lemma. *Let $p \in]1,\infty[$, $a,b \in \mathbf{R}$, and suppose $L \in [1,\infty[$ satisfies $L \geq p^*-1$. Then*

$$(L|a| - |b|)(|a|+|b|)^{p-1} \geq L^p|a|^p - |b|^p. \tag{4.6}$$

Proof. The inequality is clear when $b=0$. So suppose $a,b \in \mathbf{R}$ with $b \neq 0$.

We define $\phi : \mathbf{R} \to \mathbf{R}$ by $\phi(x) := (Lx-1)(1+x)^{p-1} - (Lx^p-1)$. Then we claim that $\phi(x) \geq 0$ whenever $x \geq 0$. It follows from this claim that

$$(L|a|-|b|)(|a|+|b|)^{p-1} - (L^p|a|^p - |b|^p) = |b|^p\phi\left(\frac{|a|}{|b|}\right) \geq 0,$$

as desired.

For the claim, we note that for $x > 0$ we have

$$\begin{aligned}\phi'(x) &= (pLx + (L-(p-1)))(1+x)^{p-2} - pLx^{p-1} \\ &= x^{p-1}\left(\left(pL + (L-(p-1))\frac{1}{x}\right)\left(\frac{1}{x}+1\right)^{p-2} - pL\right) \\ &= x^{p-1}\psi\left(\frac{1}{x}\right),\end{aligned} \tag{4.7}$$

where $\psi :]-1,\infty[\to \mathbf{R}$ is defined by

$$\psi(x) := (pL + (L-(p-1))x)(1+x)^{p-2} - pL.$$

In view of (4.7), to prove the claim it suffices to show that $\psi(x) \geq 0$ for $x \geq 0$, since this would imply that ϕ is increasing on $\mathbf{R}_+$ and thus

$$\phi(x) \geq \phi(0) = 0$$

whenever $x \geq 0$.

Since $L \geq p^* - 1 = \max(p-1, 1/(p-1))$, we find that $(p-1)L \geq 1$ and $L-(p-1) \geq 0$. Hence,

$$\psi'(x) = (p-1)(1+x)^{p-3}((p-1)L + (L-(p-1))x - 1) \geq 0$$

whenever $x \geq 0$. Thus, $\psi(x) \geq \psi(0) = 0$. The assertion follows. □

Proof of Proposition 4.3. Set $L := \|\mathscr{B}\|^p_{\mathcal{L}(L^p(\mathbf{C}))}$. Then $L \geq (p^*-1)^p \geq p^*-1 \geq 1$ by Proposition 2.45. Thus, since

$$L^p\|\partial_{\overline{z}} f\|_p \geq \|\mathscr{B}(\partial_{\overline{z}} f)\|_p^p = \|\partial_z f\|_p^p,$$

it follows from Lemma 4.4 that

$$\int_{\mathbf{C}} (L|\partial_{\overline{z}} f| - |\partial_z f(z)|)(|\partial_z f(z)| + |\partial_{\overline{z}} f(z)|)^{p-1}\,\mathrm{d}z \geq L^p\|\partial_{\overline{z}} f\|_p^p - \|\partial_z f\|_p^p \geq 0,$$

for all $f \in W^{1,p}(\mathbf{C})$. The assertion follows. □

We will use these results as a motivation to study the functional

$$E_{p,\gamma} : \mathbf{R}^{2\times 2} \to \mathbf{R}, \quad E_{p,\gamma}(A) := \gamma|A|^p - |A|^{p-2}\det A,$$

for $p \in]1,\infty[$ and $\gamma \geq |1-2/p|$. In particular, we define the following:

4.5 Definition. Let $p \in]1,\infty[$. Then we define the *Burkholder functional* as the functional

$$B_p : \mathbf{R}^{2\times 2} \to \mathbf{R}, \quad B_p(A) := \left|1 - \frac{2}{p}\right| |A|^p - |A|^{p-2}\det A.$$

♢

While it is known that such functionals are rank-one convex, it is an open problem wether they are quasiconvex. We will discuss these notions in the upcoming subsection.

4.2 Rank-one Convex and Quasiconvex functions

We provide a short introduction into the theory of calculus of variations as a motivation for the definitions of rank-one convex and quasiconvex functions. Such notions stem from the work of Morrey in 1952, see [Mo], where he discussed necessary conditions for lower semicontinuity of certain functionals.

Throughout the following example we will consider **real-valued** functions only. We will emphasize this fact by writing $;\mathbf{R}$ for our function spaces. Suppose $\Omega \subseteq \mathbf{R}^n$ is a non-empty open and bounded set so that $\overline{\Omega}$ is a C^∞ manifold with boundary. Given a function $g \in C^\infty(\partial\Omega;\mathbf{R})$, we want to find a solution $u \in C^\infty(\overline{\Omega};\mathbf{R})$ to the problem

$$\begin{cases} \Delta u = 0 \\ u|_{\partial\Omega} = g, \end{cases} \tag{4.8}$$

where $\Delta := -\sum_{j=1}^n \partial_j$ is the Laplace operator. To facilitate this, we consider a different problem where, for fixed $g \in W^{1,2}(\Omega;\mathbf{R})$, we wish to find $u \in W^{1,2}(\Omega;\mathbf{R})$ so that $\Delta u = 0$ in the distributional sense, and $u - g$ can be approximated in $W^{1,2}(\Omega;\mathbf{R})$ by a sequence in $C_c^\infty(\Omega;\mathbf{R})$, i.e., u must lie in the convex set

$$C_g := g + W_0^{1,2}(\Omega;\mathbf{R}) \subseteq W^{1,2}(\Omega;\mathbf{R}),$$

where $W_0^{1,2}(\Omega;\mathbf{R})$ is the closure of $C_c^\infty(\Omega;\mathbf{R})$ in $W^{1,2}(\Omega;\mathbf{R})$. Note that elliptic regularity implies that such a solution u must be a smooth function in Ω, see Theorem B.37. We refer to the proof of [Ni, Corollary 2.31] to see how one can solve (4.8) using the solution to the distributional problem and for a more elaborate discussion concerning these kind of boundary problems. While we note that considering the case $p = 2$ is sufficient to solve our particular problem, our arguments work just as well for any other $p \in]1,\infty[$.

We now define $F : \mathbf{R}^n \to \mathbf{R}$ by $F(x) := |x|^2$ and consider the map

$$\mathcal{E} : W^{1,2}(\Omega;\mathbf{R}) \to \mathbf{R}, \quad \mathcal{E}(u) := \int_\Omega F(Du(x))\,\mathrm{d}x,$$

where $Du = (\partial_1 u, \ldots, \partial_n u)$ is the gradient of $u \in W^{1,2}(\Omega;\mathbf{R})$. Since F is non-negative, the functional $\mathcal{E}$ is bounded from below. Our theory is motivated by the following result:

4.6 Theorem (Dirichlet's Principle). *Let $\Omega \subseteq \mathbf{R}^n$ be open and bounded and let $g \in W^{1,2}(\Omega;\mathbf{R})$. Then $u \in C_g$ satisfies $\Delta u = 0$ in the distributional sense, if and only if*

$$\mathcal{E}(u) = \inf_{v \in C_g} \mathcal{E}(v). \tag{4.9}$$

The equation $\Delta u = 0$ is called the *Euler-Lagrange equation* of $\mathcal{E}$.

Proof. Suppose $u \in C_g$ satisfies $\Delta u = 0$ in $\mathcal{D}'(\Omega)$. Let $v \in C_g$. Then $u - v = (u - g) - (v - g) \in W_0^{1,2}(\Omega;\mathbf{R})$ can be approximated in $W^{1,2}(\Omega;\mathbf{R})$ by a sequence $(\phi_j)_{j\in\mathbf{N}}$ in $C_c^\infty(\Omega;\mathbf{R})$. Hence,

$$\int_\Omega Du(x)\cdot(Du(x) - Dv(x))\,\mathrm{d}x = \lim_{j\to\infty}\int_\Omega Du(x)\cdot(D\phi_j(x))\,\mathrm{d}x = \lim_{j\to\infty}\langle \Delta u, \phi_j\rangle = 0. \tag{4.10}$$

Using the inequality

$$2|Du(x) \cdot Dv(x)| \leq 2|Du(x)||Dv(x)| \leq |Du(x)|^2 + |Dv(x)|^2,$$

we conclude from (4.10) that

$$2\mathcal{E}(u) = 2\int_\Omega |Du(x)|^2 \,\mathrm{d}x = \int_\Omega 2Du(x) \cdot Dv(x) \,\mathrm{d}x \leq \mathcal{E}(u) + \mathcal{E}(v).$$

Hence, $\mathcal{E}(u) \leq \mathcal{E}(v)$, proving that u satisfies (4.9).

For the converse, let $\phi \in C_c^\infty(\Omega; \mathbf{R})$. Then $u + t\phi \in C_g$ for all $t \in \mathbf{R}$. Setting

$$f : \mathbf{R} \to \mathbf{R}, \quad f(t) := \mathcal{E}(u + t\phi),$$

the assumption on u implies that f attains a minimum at $t = 0$. Hence,

$$0 = f'(0) = \langle Du, D\phi \rangle = \langle \Delta u, \phi \rangle.$$

We conclude that $\Delta u = 0$ in $\mathcal{D}'(\Omega)$, as desired. □

We have changed our problem of solving (4.8) to the problem of minimizing the functional $\mathcal{E}$ in C_g.

Pick a sequence $(u_j)_{j \in \mathbf{N}}$ in C_g so that

$$\lim_{j \to \infty} \mathcal{E}(u_j) = \inf_{v \in C_g} \mathcal{E}(v) =: \kappa. \tag{4.11}$$

Recall that the Poincaré inequality states that, since Ω is bounded, there is a constant $c \in \mathbf{R}_+$ so that

$$\|v\|^2_{W^{1,2}(\Omega)} \leq c \int_\Omega |Dv(x)|^2 \,\mathrm{d}x = c\mathcal{E}(v)$$

for all $v \in W_0^{1,2}(\Omega; \mathbf{R})$. By (4.11), this implies that the sequence $(u_j)_{j \in \mathbf{N}}$ is bounded in $W^{1,2}(\Omega; \mathbf{R})$. Thus, since $W^{1,2}(\Omega; \mathbf{R})$ is a reflexive Banach space, it follows from the Banach-Alaoglu Theorem that there is a weakly convergent subsequence $(u_{j_k})_{k \in \mathbf{N}}$ of $(u_j)_{j \in \mathbf{N}}$ with limit $u \in W^{1,2}(\Omega; \mathbf{R})$. Since C_g is a closed convex subset of $W^{1,2}(\Omega; \mathbf{R})$, it follows that C_g is also weakly closed in $W^{1,2}(\Omega; \mathbf{R})$. Thus, we must have $u \in C_g$.

It now remains to show that $\mathcal{E}(u) = \kappa$. Since $\kappa \leq \mathcal{E}(u)$, we only need the converse inequality. A sufficient condition on $\mathcal{E}$ is that

$$\mathcal{E}(v) \leq \liminf_{j \to \infty} \mathcal{E}(v_j) \tag{4.12}$$

for every sequence $(v_j)_{j \in \mathbf{N}}$ in $W^{1,2}(\Omega; \mathbf{R})$ that weakly converges to $v \in W^{1,2}(\Omega; \mathbf{R})$, since then

$$\mathcal{E}(u) \leq \liminf_{k \to \infty} \mathcal{E}(u_{j_k}) = \kappa.$$

The property (4.12) is usually referred to as *weak lower semicontinuity* of $\mathcal{E}$.

To see that $\mathcal{E}$ is weakly lower semicontinuous, we can appeal to a general result in Banach spaces.

4.7 Lemma. *Let X be a topological vector space with seminorm $\|\cdot\|$ and let X^* denote its dual space. If a sequence $(x_j)_{j\in\mathbf{N}}$ in X converges weakly to $x \in X$, then*

$$\|x\|^2 \leq \liminf_{j\to\infty} \|x_j\|^2.$$

Proof. By Hahn-Banach we can pick $x^* \in X^*$ so that $\langle x^*, x\rangle = \|x\|$ and $|\langle x^*, y\rangle| \leq \|y\|$ for all $y \in X$. Then the inequality

$$|\langle x^*, x_j\rangle|^2 \leq \|x_j\|^2$$

for all $j \in \mathbf{N}$ implies that

$$\|x\|^2 = \lim_{j\to\infty} |\langle x^*, x_j\rangle|^2 \leq \liminf_{j\to\infty} \|x_j\|^2.$$

The assertion follows. □

Since $u \mapsto (\mathcal{E}(u))^{\frac{1}{2}}$ defines a seminorm on $W^{1,2}(\Omega;\mathbf{R})$, we may immediately deduce from the lemma that

$$\mathcal{E}(v) \leq \liminf_{j\to\infty} \mathcal{E}(v_j)$$

for any sequence $(v_j)_{j\in\mathbf{N}}$ in $W^{1,2}(\Omega;\mathbf{R})$ that converges weakly to $v \in W^{1,2}(\Omega;\mathbf{R})$, proving the desired result.

The general cases of the situation described in our example are well understood in the sense that we have the following result:

4.8 Theorem. *Let $\Omega \subseteq \mathbf{R}^n$ be open and bounded. Let $F : \mathbf{R}^n \to \mathbf{R}$ and let $\mathcal{E} : W^{1,2}(\Omega;\mathbf{R}) \to \mathbf{R}$ be defined by*

$$\mathcal{E}(u) := \int_\Omega F(Du(x))\,\mathrm{d}x.$$

Then $\mathcal{E}$ is weakly lower semicontinuous if and only if F is convex.

See [Ev2, Theorem 2.2.1] for a proof. While this theorem deals with real-valued functions, one also wishes to consider the vector-valued case, i.e., functions taking values in $\mathbf{R}^m$.

We denote by $\mathbf{R}^{m\times n}$ the set of $m \times n$ matrices with real coefficients. We equip it with the operator norm

$$|A| := \sup_{|h|=1} |Ah|.$$

For an open and bounded $\Omega \subseteq \mathbf{R}^n$ we denote by $W^{1,1}_{loc}(\Omega;\mathbf{R}^m)$ the space of those $u = (u_1,\ldots,u_m)$ with $u_j \in W^{1,1}_{loc}(\Omega;\mathbf{R})$ for all $j \in \{1,\ldots,m\}$. For any $u \in W^{1,1}_{loc}(\Omega;\mathbf{R}^m)$ we may then define the total derivative Du by the (generalized) Jacobian matrix

$$Du(x) := \begin{pmatrix} \partial_1 u_1(x) & \cdots & \partial_n u_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 u_m(x) & \cdots & \partial_n u_m(x) \end{pmatrix} \in \mathbf{R}^{m\times n}$$

for a.e. $x \in \Omega$.

4.9 Definition. Let $F : \mathbf{R}^{m\times n} \to \mathbf{R}$ be continuous. We say that F is *quasiconvex at* $A \in \mathbf{R}^{m\times n}$ if

$$\int_\Omega F(A + D\phi(x))\,\mathrm{d}x \geq \int_\Omega F(A)\,\mathrm{d}x = |\Omega|F(A)$$

for all open and bounded $\Omega \subseteq \mathbf{R}^n$ and all $\phi \in C_c^\infty(\Omega; \mathbf{R}^m)$. We say that F is *quasiconvex* if F is quasiconvex at A for all $A \in \mathbf{R}^{m\times n}$. ♢

This definition is motivated through Morrey's search for necessary conditions for lower semicontinuity of certain functionals. Instead of perturbations by compactly supported smooth functions, he considered perturbations by Lipschitz continuous functions with vanishing boundary conditions. These notions of quasiconvexity are the same as can be shown by a density argument, see Proposition 4.12 below. We remark that any Lipschitz continuous function is differentiable almost everywhere by Rademacher's Theorem. Furthermore, Morrey considered lower semicontinuity with respect to convergence of functions in the space of Lipschitz continuous functions rather than weak lower semicontinuity in our sense:

4.10 Definition. Let $\Omega \subseteq \mathbf{R}^n$ be non-empty open and bounded and let $p \in]1, \infty[$. We say that $\mathcal{E} : W^{1,p}(\Omega; \mathbf{R}^m) \to \mathbf{R}$ is *weakly lower semicontinuous* if

$$\mathcal{E}(v) \leq \liminf_{j\to\infty} \mathcal{E}(v_j)$$

for every weakly convergent sequence $(v_j)_{j\in\mathbf{N}}$ in $W^{1,p}(\Omega; \mathbf{R}^m)$ with limit $v \in W^{1,p}(\Omega; \mathbf{R}^m)$. ♢

We note that Conjecture 4.2 is the statement that the Burkholder functional is quasiconvex at 0. While we are now considering a more modern setting, the ideas used remain the same.

4.11 Theorem. *Let $\Omega \subseteq \mathbf{R}^n$ be a non-empty open and bounded set and let $p \in]1, \infty[$. Let $F : \mathbf{R}^{m\times n} \to \mathbf{R}$ be continuous and let $\mathcal{E} : W^{1,p}(\Omega; \mathbf{R}^m) \to \mathbf{R}$ be defined by*

$$\mathcal{E}(u) := \int_\Omega F(Du(x))\,\mathrm{d}x.$$

If $\mathcal{E}$ is weakly lower semicontinuous, then F is quasiconvex.

Conversely, if F is quasiconvex and there is some $c \in \mathbf{R}_+$ so that

$$0 \leq F(A) \leq c(1 + |A|^p)$$

for all $A \in \mathbf{R}^{m\times n}$, then $\mathcal{E}$ is weakly lower semicontinuous.

A proof can be found in [Ev2, Theorem 3.2.1].

Now it may seem strange that the notion of weak lower semicontinuity considers only a single open bounded set $\Omega \subseteq \mathbf{R}^n$, while in the definition of quasiconvexity we require an estimate for all open and bounded $\Omega \subseteq \mathbf{R}^n$. This is justified by the equivalence of (i) and (iv) in the following result:

4.12 Proposition. *Let $F : \mathbf{R}^{m\times n} \to \mathbf{R}$ be continuous and let $A \in \mathbf{R}^{m\times n}$. Then the following are equivalent:*

(i) F *is quasiconvex at* A*;*

(ii) *for all* $\phi \in C_c^\infty(\mathbf{R}^n;\mathbf{R}^m)$ *we have*

$$\int_{\mathbf{R}^n} (F(A + D\phi(x)) - F(A))\,\mathrm{d}x \geq 0;$$

(iii) *for all bounded open sets* $\Omega \subseteq \mathbf{R}^n$ *and all Lipschitz continuous functions* $f : \overline{\Omega} \to \mathbf{R}^m$ *satisfying* $f|_{\partial\Omega} = 0$*, we have*

$$\int_\Omega F(A + Df(x))\,\mathrm{d}x \geq |\Omega| F(A);$$

(iv) *there is a non-empty bounded open set* $\Omega \subseteq \mathbf{R}^n$ *so that for all* $\phi \in C_c^\infty(\Omega;\mathbf{R}^m)$ *we have*

$$\int_\Omega F(A + D\phi(x))\,\mathrm{d}x \geq |\Omega| F(A);$$

(v) *for each* $\psi \in C^\infty(\mathbf{R}^n;\mathbf{R}^m)$ *which is* 1*-periodic in the sense that* $\psi(x+k) = \psi(x)$ *for all* $k \in \mathbf{Z}^n$ *and* $x \in \mathbf{R}^n$*, we have*

$$\int_{[0,1]^n} F(A + D\psi(x))\,\mathrm{d}x \geq F(A).$$

Proof. For (i)⇒(ii), we pick $\phi \in C_c^\infty(\mathbf{R}^n;\mathbf{R}^m)$. Then there is some open ball $\Omega \subseteq \mathbf{R}^n$ so that $\operatorname{supp}\phi \subseteq \Omega$. Hence, we may view ϕ as an element of $C_c^\infty(\Omega;\mathbf{R}^m)$ and

$$\begin{aligned}\int_{\mathbf{R}^n} (F(A + D\phi(x)) - F(A))\,\mathrm{d}x &= \int_\Omega (F(A + D\phi(x)) - F(A))\,\mathrm{d}x \\ &= \int_\Omega F(A + D\phi(x))\,\mathrm{d}x - |\Omega| F(A) \geq 0.\end{aligned}$$

The assertion (ii)⇒(i) follows by noting that any $\phi \in C_c^\infty(\Omega;\mathbf{R}^m)$ for some open and bounded $\Omega \subseteq \mathbf{R}^n$ extends by 0 to an element of $C_c^\infty(\mathbf{R}^n;\mathbf{R}^m)$.

For (i)⇒(iii), we pick any open and bounded $\Omega \subseteq \mathbf{R}^n$ and a Lipschitz continuous function $f : \overline{\Omega} \to \mathbf{R}^m$ satisfying $f|_{\partial\Omega} = 0$. Let $(\phi_\varepsilon)_{\varepsilon\in\mathbf{R}_+}$ denote the standard mollifier, see Definition A.10. We may extend f continuously to $\mathbf{R}^n$ by declaring that it vanishes outside of Ω. Performing componentwise convolution, we may consider $f * \phi_\varepsilon \in C^\infty(\mathbf{R}^n;\mathbf{R}^m)$ for $\varepsilon \in \mathbf{R}_+$. By Lemma A.12 we have

$$\operatorname{supp}(f * \phi_\varepsilon) \subseteq \overline{\Omega} + \overline{B}_\varepsilon(0),$$

where $\overline{B}_\varepsilon(x)$ denotes the closed ball around $x \in \mathbf{R}^n$ of radius $\varepsilon \in \mathbf{R}_+$. Since this sum of sets is compact and shrinks as ε does, we can pick an open and bounded $\Omega' \subseteq \mathbf{R}^n$ that contains Ω and so that $\operatorname{supp}(f * \phi_\varepsilon) \subseteq \Omega'$ for all $\varepsilon \in \mathbf{R}_+$ small enough. For such $\varepsilon \in \mathbf{R}_+$ we may view $f * \phi_\varepsilon$ as an element of $C_c^\infty(\Omega';\mathbf{R}^m)$, which we denote by f_ε. This implies that

$$\int_{\Omega'} F(A + Df_\varepsilon(x))\,\mathrm{d}x \geq |\Omega'| F(A). \tag{4.13}$$

Since $\partial_j(f_k * \phi_\varepsilon) = \partial_j f_k * \phi_\varepsilon$ for all $j \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, m\}$, where $f = (f_1, \ldots, f_m)$, we claim that $D(f_\varepsilon)(x) \to Df(x)$ in $\mathbf{R}^{m\times n}$ as $\varepsilon \downarrow 0$ for a.e. $x \in \Omega'$. Here we note that f is almost everywhere differentiable in Ω' by Rademacher's Theorem, and the derivatives of f are bounded by its Lipschitz constant.

For the claim, we write $g := \partial_j f_k \in L^\infty(\mathbf{R}^n) \subseteq L^1_{loc}(\mathbf{R}^n)$. Then, for some $c \in \mathbf{R}_+$,

$$\begin{aligned} |(g * \phi_\varepsilon)(x) - g(x)| &= \left| \frac{1}{\varepsilon^n} \int_{\overline{B}_\varepsilon(x)} \phi\left(\frac{x-y}{\varepsilon}\right)(g(y) - g(x))\,\mathrm{d}y \right| \\ &\le \frac{c}{|\overline{B}_\varepsilon(y)|} \int_{\overline{B}_\varepsilon(y)} |g(y) - g(x)|\,\mathrm{d}y \to 0 \quad \text{as} \quad \varepsilon \downarrow 0 \end{aligned}$$

for a.e. $x \in \mathbf{R}^n$ by Lebesgue's Differentiation Theorem. This proves the claim.

By continuity of F we may apply Lebesgue's Dominated Convergence Theorem in (4.13) to conclude that

$$\begin{aligned} \int_\Omega F(A + Df(x))\,\mathrm{d}x &= \int_{\Omega'} F(A + Df(x))\,\mathrm{d}x - |\Omega' \backslash \Omega| F(A) \\ &\ge |\Omega'| F(A) - |\Omega' \backslash \Omega| F(A) = |\Omega| F(A), \end{aligned}$$

as desired. For (iii)⇒(i), we note that for all open and bounded $\Omega \subseteq \mathbf{R}^n$, any $\phi \in C_c^\infty(\Omega; \mathbf{R}^m)$ is Lipschitz continuous with vanishing boundary conditions. The result follows.

The implication (i)⇒(iv) is immediate. For the converse, we assume (iv) and let $\Omega' \subseteq \mathbf{R}^n$ be any non-empty open bounded set. Now consider the collection

$$\mathcal{V} := \{a + t\overline{\Omega'} \mid a \in \mathbf{R}^n,\, t \in \mathbf{R}_+\}.$$

As this is a Vitali covering of Ω, we may appeal to Vitali's Covering Theorem for the Lebesgue measure to obtain an at most countable disjoint family $\{a_j + t_j\overline{\Omega'}\}_j \subseteq \mathcal{V}$ of subsets of Ω so that

$$\left| \Omega \backslash \bigcup_j a_j + t_j \overline{\Omega'} \right| = 0.$$

This means that

$$\sum_j t_j^n |\Omega'| = \sum_j \left| a_j + t_j \overline{\Omega'} \right| = \left| \bigcup_j a_j + t_j \overline{\Omega'} \right| = |\Omega|. \tag{4.14}$$

Now pick $\phi \in C_c^\infty(\Omega'; \mathbf{R}^m)$ and define $\psi \in C_c^\infty(\Omega; \mathbf{R}^m)$ by

$$\psi(x) := \begin{cases} t_j \phi\left(\dfrac{x - a_j}{t_j}\right) & \text{if } x \in a_j + t_j\Omega' \\ 0 & \text{otherwise.} \end{cases}$$

Then, by (4.14),

$$\begin{aligned} |\Omega| F(A) &\le \int_\Omega F(A + D\psi(x))\,\mathrm{d}x = \sum_j \int_{a_j + t_j\Omega'} F\left(A + D\phi\left(\frac{x - a_j}{t_j}\right)\right)\mathrm{d}x \\ &= \left(\sum_j t_j^n\right) \int_{\Omega'} F(A + D\phi(x))\,\mathrm{d}x = \frac{|\Omega|}{|\Omega'|} \int_{\Omega'} F(A + D\phi(x))\,\mathrm{d}x, \end{aligned}$$

so that

$$\int_{\Omega'} F(A + D\phi(x))\,\mathrm{d}x \geq |\Omega'|F(A),$$

as desired.

For (v)⇒(iv) we note that any $\phi \in C_c^\infty(]0,1[^n;\mathbf{R}^m)$ can be extended 1-periodically to a smooth map $\psi \in C^\infty(\mathbf{R}^n;\mathbf{R}^m)$. Then

$$\int_{]0,1[^n} F(A + D\phi(x))\,\mathrm{d}x = \int_{[0,1]^n} F(A + D\psi(x))\,\mathrm{d}x \geq F(A) = |]0,1[^n|F(A),$$

proving (iv) with $\Omega =]0,1[^n$.

Finally, for (i)⇒(v) we refer to [Da, Proposition 5.13]. This concludes the proof.

□

4.13 Remark. We note that for an open and bounded set $\Omega \subseteq \mathbf{R}^n$ with C^1-boundary, the space of Lipschitz continuous functions in Ω coincides with the space $W^{1,\infty}(\Omega)$, see [Ev, 5.8, Theorem 5]. Thus, in view of characterization (iii) in the proposition, it makes sense that the notion of quasiconvexity we are considering here is sometimes called $W^{1,\infty}$-quasiconvexity as apposed to the general notion of $W^{1,p}$-quasiconvexity, introduced in [BM], that deals with perturbations in the space $W_0^{1,p}(\Omega;\mathbf{R}^m)$ for $p \in [1,\infty]$.

We have included the characterization (v) in the proposition since this was famously used by Šverák in [Šv] to construct a function $F : \mathbf{R}^{3\times 2} \to \mathbf{R}$ that is rank-one convex, see Definition 4.17 below, while it is not quasiconvex. For a further discussion of this example, we refer to Theorem 4.27 below. ♢

Having given our definition of quasiconvexity, it's reasonable to check that quasiconvexity is indeed a consequence of convexity.

4.14 Proposition. *If $F : \mathbf{R}^{m\times n} \to \mathbf{R}$ is continuous and convex, then F is quasiconvex. Conversely, if $n = 1$ or $m = 1$ and F is quasiconvex, then F is convex.*

Proof. Let $\Omega \subseteq \mathbf{R}^n$ be a non-empty open and bounded set, let $A \in \mathbf{R}^{m\times n}$, and let $\phi \in C_c^\infty(\Omega;\mathbf{R}^m)$. Considering componentwise integration, we note that

$$\int_\Omega D\phi(x)\,\mathrm{d}x = 0$$

by the Fundamental Theorem of Calculus. Then it follows from Jensen's inequality that

$$\frac{1}{|\Omega|}\int_\Omega F(A + D\phi(x))\,\mathrm{d}x \geq F\left(\frac{1}{|\Omega|}\int_\Omega (A + D\phi(x))\,\mathrm{d}x\right) = F(A).$$

This proves the first assertion.

Now suppose $n = 1$ and suppose F is quasiconvex. We identify $\mathbf{R}^{1\times m}$ with $\mathbf{R}^m$. Fix $a, b \in \mathbf{R}^m$ and let $t \in]0,1[$. We define $c := ta + (1-t)b$ and define

$$f : \mathbf{R} \to \mathbf{R}^m, \quad f(x) := a\chi_{]0,t[}(x) + b\chi_{]t,1[}(x),$$

where $\chi_{]r,s[}$ denotes the indicator function of the interval $]r,s[$, $r,s\in\mathbf{R}$, $r<s$. Then

$$\phi:[0,1]\to\mathbf{R}^m,\quad \phi(y):=-yc+\int_0^y f(x)\,\mathrm{d}x$$

defines a Lipschitz continuous map with derivative $\phi'(x)=-c+f(x)$ for a.e. $x\in[0,1]$ and

$$\phi(1)=-c+\int_0^t a\,\mathrm{d}x+\int_t^1 b\,\mathrm{d}x=-c+c=0=\phi(0).$$

Thus, ϕ is a Lipschitz continuous map with vanishing boundary conditions. Hence, quasiconvexity of F (in the sense of Morrey) implies that

$$F(ta+(1-t)b)=F(c)\le\int_0^1 F(c+\phi'(x))\,\mathrm{d}x=\int_0^t F(a)\,\mathrm{d}x+\int_t^1 F(b)\,\mathrm{d}x=tF(a)+(1-t)F(b).$$

This proves that F is convex, as desired. The case $m=1$ is analogous. The assertion follows. □

4.15 Remark. Since we are interested in the quasiconvexity of the Burkhulder functional, it would make sense to check if the Burkholder functional is convex. As a simple example shows, this is unfortunately not the case. Let $p\in]1,\infty[$ and denote by $I\in\mathbf{R}^{2\times 2}$ the identity matrix. Then we note that

$$B_p(I)=\left|1-\frac{2}{p}\right|-1=1-\frac{2}{p^*}-1=-\frac{2}{p^*}<0.$$

Hence,

$$B_p\left(\frac{1}{2}0+\frac{1}{2}I\right)=\frac{1}{2^p}B_p(I)>\frac{1}{2}B_p(I)=\frac{1}{2}B_p(0)+\frac{1}{2}B_p(I).$$

Thus, B_p is indeed not convex. ♢

Now, rather than considering lower semicontinuity, let us consider a more direct approach to minimizing a functional

$$\mathcal{E}(u)=\int_\Omega F(Du(x))\,\mathrm{d}x,\quad u\in W^{1,p}(\Omega;\mathbf{R}^m)$$

for $p\in]1,\infty[$, a given open and bounded $\Omega\subseteq\mathbf{R}^n$, and a given $F:\mathbf{R}^{m\times n}\to\mathbf{R}$. For a given $g\in W^{1,p}(\Omega;\mathbf{R}^m)$ we again consider the functions $u\in W^{1,p}(\Omega;\mathbf{R}^m)$ with the boundary condition $u-g\in W_0^{1,p}(\Omega;\mathbf{R}^m)$, where $W_0^{1,p}(\Omega;\mathbf{R}^m)$ denotes the closure of $C_c^\infty(\Omega;\mathbf{R}^m)$ in $W^{1,p}(\Omega;\mathbf{R}^m)$. Suppose that $v\in W^{1,p}(\Omega;\mathbf{R}^m)$ minimizes $\mathcal{E}$ in this class of functions satisfying the boundary condition. Now pick any Lipschitz continuous function ψ with compact support in Ω and define

$$f:\mathbf{R}\to\mathbf{R},\quad f(t):=\mathcal{E}(v+t\psi).$$

Since $v+t\psi$ satisfies the desired boundary conditions for all $t\in\mathbf{R}$, the function f must be minimized at 0. Assuming that F is sufficiently regular, this means that the second derivative of f at 0 must be non-negative. Denoting by $D^2F(A):\mathbf{R}^{n\times m}\times\mathbf{R}^{n\times m}\to\mathbf{R}$ the Hessian of F at $A\in\mathbf{R}^{n\times m}$, we have

$$0\le f''(0)=\int_\Omega D^2F(Dv(x))(D\psi(x),D\psi(x))\,\mathrm{d}x. \tag{4.15}$$

Now let $\rho : \mathbf{R} \to \mathbf{R}$ be the 2-periodic sawtooth function, i.e., the 2-periodic extension of the function that is equal to t on $[0,1]$ and equal to $2-t$ on $[0,1]$. Then ρ is a Lipschitz continuous function that satisfies $\rho'(t)^2 = 1$ for a.e. $t \in \mathbf{R}$. Pick any pair of vectors $\eta \in \mathbf{R}^m$, $\xi \in \mathbf{R}^n$, and let $\varepsilon \in \mathbf{R}_+$. Letting $\phi \in C_c^\infty(\Omega; \mathbf{R})$ be arbitrary, we define

$$\psi_\varepsilon : \Omega \to \mathbf{R}^m, \quad \psi_\varepsilon(x) := \varepsilon \phi(x) \rho\left(\frac{\xi \cdot x}{\varepsilon}\right) \eta.$$

Then ψ_ε is a Lipschitz continuous function of compact support in Ω. Picking $i \in \{1, \ldots, n\}$, $k \in \{1, \ldots, m\}$ we note that if ψ_ε^k denotes the k-th component of ψ_ε, then

$$\partial_i \psi_\varepsilon^k(x) = \varepsilon \partial_i \phi(x) \rho\left(\frac{\xi \cdot x}{\varepsilon}\right) \eta_k + \phi(x) \rho'\left(\frac{\xi \cdot x}{\varepsilon}\right) \xi_i \eta_k.$$

This means that for any pair $i, j \in \{1, \ldots, n\}$ and any pair $k, l \in \{1, \ldots, m\}$ we have

$$\begin{aligned} \partial_i \psi_\varepsilon^k(x) \partial_j \psi_\varepsilon^l(x) &= \varepsilon^2 \partial_i \phi(x) \partial_j \phi(x) \rho\left(\frac{\xi \cdot x}{\varepsilon}\right)^2 \eta_k \eta_l + \phi(x)^2 \rho'\left(\frac{\xi \cdot x}{\varepsilon}\right)^2 \xi_i \eta_k \xi_j \eta_l \\ &\quad + \varepsilon \left(\phi(x) \partial_i \phi(x) \rho\left(\frac{\xi \cdot x}{\varepsilon}\right) \rho'\left(\frac{\xi \cdot x}{\varepsilon}\right) \eta_k \xi_j \eta_l + \phi(x) \partial_j \phi(x) \rho\left(\frac{\xi \cdot x}{\varepsilon}\right) \rho'\left(\frac{\xi \cdot x}{\varepsilon}\right) \xi_i \eta_k \eta_l \right) \end{aligned}$$

for a.e. $x \in \Omega$. Since ρ and ρ' are bounded, while $\rho'(t)^2 = 1$ for a.e. $t \in \mathbf{R}$, we conclude that

$$\partial_i \psi_\varepsilon^k(x) \partial_j \psi_\varepsilon^l(x) \to \phi(x)^2 \xi_i \eta_k \xi_j \eta_l \quad \text{as} \quad \varepsilon \downarrow 0$$

for a.e. $x \in \Omega$.

Since the Hessian $D^2F(Dv(x))$, $x \in \Omega$, is bilinear, replacing ψ by ψ_ε in (4.15) and letting $\varepsilon \downarrow 0$ implies that

$$\int_\Omega D^2F(Dv(x))(\eta \otimes \xi, \eta \otimes \xi) \phi(x)^2 \, \mathrm{d}x \geq 0$$

for all $\phi \in C_c^\infty(\Omega; \mathbf{R})$, where $\eta \otimes \xi \in \mathbf{R}^{m \times n}$ denotes the matrix with entry $\eta_k \xi_i$ on the (k,i)-th position. Since $\phi \in C_c^\infty(\Omega; \mathbf{R})$ is arbitrary, we conclude that

$$D^2F(Dv(x))(\eta \otimes \xi, \eta \otimes \xi) \geq 0$$

for a.e. $x \in \Omega$. Since $\eta \otimes \xi$ is a typical rank-one matrix, it is therefore not unreasonable to assume that F satisfies the so-called *Legendre-Hadamard condition*

$$D^2F(A)(X,X) \geq 0 \quad \text{for all } A, X \in \mathbf{R}^{m \times n} \text{ with } \operatorname{rank} X = 1. \tag{4.16}$$

We recall the following basic result on convex functions on the real line:

4.16 Lemma. *Let $U \subseteq \mathbf{R}$ be an open interval and suppose $f : U \to \mathbf{R}$ is twice differentiable. If $f''(t) \geq 0$ for all $t \in U$, then f is convex in U.*

Proof. First note that the condition $f''(t) \geq 0$ for all $t \in U$ implies that f' is increasing in U. Now let $x, y \in U$, $x < y$, and $t \in [0,1]$. We set $p := tx + (1-t)y$. Then we have

$$\begin{aligned} tf(x) + (1-t)f(y) - f(p) &= tf(p) + t\int_p^x f'(s)\,\mathrm{d}s + (1-t)f(p) + (1-t)\int_p^y f'(s)\,\mathrm{d}s - f(p) \\ &= -t\int_x^p f'(s)\,\mathrm{d}s + (1-t)\int_p^y f'(s)\,\mathrm{d}s \\ &\geq -t(p-x)f'(p) + (1-t)(y-p)f'(p) \\ &= t(1-t)(x-y)f'(p) + t(1-t)(y-x)f'(p) = 0. \end{aligned}$$

The assertion follows. □

The condition (4.16) means that the function $t \mapsto F(A+tX)$ is convex in $\mathbf{R}$ for every $A, X \in \mathbf{R}^{m\times n}$ with X a rank-one matrix. Heuristically, this means that this so called rank-one convexity is a natural condition to impose on F when minimizing $\mathcal{E}$. Note also that in the case $m = 1$ our considerations reduce to the convexity condition from Theorem 4.8.

4.17 Definition. Let $F : \mathbf{R}^{m\times n} \to \mathbf{R}$. We say that F is *rank-one convex at* $A \in \mathbf{R}^{m\times n}$ if $t \mapsto F(A+tX)$ is a convex function in $\mathbf{R}$ for every matrix $X \in \mathbf{R}^{m\times n}$ of rank one. We call F *rank-one convex* if it is rank-one convex at A for all $A \in \mathbf{R}^{m\times n}$. ♢

Generally, checking if a function is rank-one convex is easier than checking if a function is quasiconvex, since pointwise estimates are simpler to establish than integral estimates.

Since we presented quasiconvexity as a characterization of weak lower semicontinuity of the associated integral functional while we presented rank-one convexity as a necessary condition for having minimizers of the associated integral functional, the following result is not surprising:

4.18 Proposition. *Let $F : \mathbf{R}^{m\times n} \to \mathbf{R}$ be quasiconvex at $A \in \mathbf{R}^{m\times n}$. Then F is rank-one convex at A.*

For a proof we refer to [Mo, Theorem 4.1 & 4.2]. Thus, if the Burkholder functional satisfies the conjectured quasiconvexity property, then it must certainly be rank-one convex. We will now prove that this is indeed the case.

4.19 Theorem (Burkholder)**.** *Let $p \in]1,\infty[$, let $\gamma \in \mathbf{R}$ with $\gamma \geq |1-2/p|$, and let $A, X \in \mathbf{R}^{2\times 2}$ with $\det X \leq 0$. Then the function $f_\gamma : \mathbf{R} \to \mathbf{R}$,*

$$f_\gamma(t) := E_{p,\gamma}(A+tX) = \gamma|A+tX|^p - |A+tX|^{p-2}\det(A+tX)$$

is convex. In particular, the functionals $E_{p,\gamma}$, including the Burkholder functional B_p, are rank-one convex.

Before we turn to the proof of this result, we mention a generalization that was proven by Iwaniec in [Iw2], which emphasizes the critical nature of the Burkholder functional.

4.20 Theorem (Iwaniec, 2002)**.** *Let $p \in]1,\infty[$ and let $\gamma \in \mathbf{R}$. Then the functionals $E^{\pm}_{p,\gamma,n} : \mathbf{R}^{n\times n} \to \mathbf{R}$ defined by*

$$E^{\pm}_{p,\gamma,n}(A) := \gamma|A|^p \pm |A|^{p-n}\det(A)$$

are rank-one convex for $\gamma \geq |1-n/p|$ and not rank-one convex for $\gamma < |1-n/p|$.

Now, while the task of proving Theorem 4.19 seems daunting, the strategy for the proof we give is straightforward. First we reduce to the cases where A is given by a simple matrix. Then we conclude the proof by applying the second derivative test for convexity.

4.21 Lemma. *Let $p \in]1, \infty[$ and let $A, X \in \mathbf{R}^{2\times 2}$. Then the map $t \mapsto |A + tX|^p$ is convex in $\mathbf{R}$.*

Proof. First we will show that the map $r : [0, \infty[\to \mathbf{R}$, $r(s) := s^p$ is convex and increasing in $[0, \infty[$. We have $r''(s) = p(p-1)s^{p-2} \geq 0$ for all $s \in]0, \infty[$, proving convexity of r in $]0, \infty[$ by Lemma 4.16. If $x = 0$, $y > 0$, $t \in [0, 1]$, we have

$$(tx + (1-t)y)^p = (1-t)^p y^p \leq (1-t)y^p = tx^p + (1-t)y^p,$$

establishing convexity of r at 0 so that r is indeed convex. To see that it's increasing we note that $r'(s) = ps^{p-1} \geq 0$ for all $s \in [0, \infty[$.

Next, for all $x, y \in \mathbf{R}$, $t \in [0, 1]$ we have

$$|A + (tx + (1-t)y)X| = |tA + txX + (1-t)A + (1-t)yX| \leq t|A + xX| + (1-t)|A + yX|$$

so that

$$\begin{aligned} |A + (tx + (1-t)yX)|^p &= r(|A + (tx + (1-t)yX)|) \leq r(t|A + xX| + (1-t)|A + yX|) \\ &\leq tr(|A + xX|) + (1-t)r(|A + yX|) = t|A + xX|^p + (1-t)|A + yX|^p. \end{aligned}$$

The assertion follows. □

4.22 Lemma. *Let $A \in \mathbf{R}^{2\times 2}$ with $|A| = 1$, $|\det A| \neq 1$. Then there exist rotation matrices $O_1, O_2 \in \mathbf{R}^{2\times 2}$ so that*

$$O_1 A O_2 = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}$$

for some $\lambda \in \mathbf{R}$ with $|\lambda| < 1$.

Proof. Using the singular value decomposition of 2×2 matrices we can find rotation matrices $O_1, O_2 \in \mathbf{R}^{2\times 2}$ so that $O_1 A O_2$ is diagonal, where the diagonal entries are given by roots of the eigenvalues of $A^t A$. Denote the diagonal entries by $\lambda, \mu \in \mathbf{R}_{\geq 0}$. We claim that one of these must have absolute value 1 and the other must have absolute value strictly smaller than 1. Since multiplication by rotation matrices leaves the operator norm invariant, we have, by Lemma 3.7, that

$$1 = |A| = |O_1 A O_2| = \frac{|\lambda + \mu| + |\lambda - \mu|}{2} = \frac{||\lambda| + |\mu|| + ||\lambda| - |\mu||}{2} = \max(|\lambda|, |\mu|).$$

Assume that $|\mu| = 1$. Then $|\lambda| \leq 1$, but since

$$1 \neq |\det A| = |\det O_1 \det A \det O_2| = |\det O_1 A O_2| = |\lambda\mu| = |\lambda|,$$

we may conclude that $|\lambda| < 1$, as desired.

Then, if necessary, by altering O_1 and O_2 by applying the rotation matrix $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ from the left and the rotation matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ from the right to change the position of the diagonal entries and by multiplying by minus the identity if $\mu = -1$, we may assume that $O_1 A O_2$ is of the desired form. This proves the assertion. □

4.23 Lemma. *Let* $A, X \in \mathbf{R}^{2\times 2}$ *and*

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}, \quad |\lambda| < 1, \quad X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Moreover, assume that $|A + tX|^2 \neq |\det(A + tX)|$ *for all* $t \in \mathbf{R}$*. Define* $h : \mathbf{R} \to \mathbf{R}$ *by* $h(t) := |A + tX|$*. Then* $h \in C^\infty(\mathbf{R}; \mathbf{R})$*,* h *is nowhere vanishing, and*

$$h'(0) = d, \quad h''(0) = \frac{b^2 + c^2 + 2\lambda bc}{1 - \lambda^2}.$$

Proof. By Lemma 3.7 we have

$$h(t) = |z(t)| + |w(t)|, \quad 2z(t) = \lambda + 1 + t(a + d + i(c - b)), \quad 2w(t) = \lambda - 1 + t(a - d + i(c + b))$$

and

$$|\det(A + tX)| = |A + tX|||z(t)| - |w(t)||. \tag{4.17}$$

We claim that z and w are nowhere vanishing, so that h is a nowhere vanishing smooth function. By (4.17), the assumption that $|A + tX|^2 \neq |\det(A + tX)|$ for all $t \in \mathbf{R}$ implies that

$$||z(t)| - |w(t)|| \neq |A + tX| = |z(t)| + |w(t)| \quad \text{for all } t \in \mathbf{R}.$$

But since $||z(t)| - |w(t)|| \leq |z(t)| + |w(t)|$, this means that $||z(t)| - |w(t)|| < |z(t)| + |w(t)|$ for all $t \in \mathbf{R}$. If $z(t) = 0$ for some $t \in \mathbf{R}$, this means that we must have $|w(t)| < |w(t)|$, which is absurd. Similarly, we cannot have $w(t) = 0$ for any $t \in \mathbf{R}$. The claim follows.

We may compute

$$\begin{aligned} h'(t) &= \frac{\mathrm{Re}(z(t)\overline{z'(t)})}{|z(t)|} + \frac{\mathrm{Re}(w(t)\overline{w'(t)})}{|w(t)|} \\ &= \frac{(\lambda + 1 + t(a + d))(a + d) + t(c - b)^2}{4|z(t)|} + \frac{(\lambda - 1 + t(a - d))(a - d) + t(c + b)^2}{4|w(t)|}, \end{aligned}$$

from which we conclude that

$$h'(0) = \frac{(\lambda + 1)(a + d)}{2|\lambda + 1|} + \frac{(\lambda - 1)(a - d)}{2|\lambda - 1|} = \frac{a + d - (a - d)}{2} = d.$$

Finally, we compute

$$\begin{aligned} h''(0) &= \frac{|z(0)|((a + d)^2 + (c - b)^2) - (a + d)^2 \frac{\lambda + 1}{2}}{4|z(0)|^2} + \frac{|w(0)|((a - d)^2 + (c + b)^2) - (a - d)^2 \frac{1 - \lambda}{2}}{4|w(0)|^2} \\ &= \frac{(c - b)^2}{2(1 + \lambda)} + \frac{(c + b)^2}{2(1 - \lambda)} = \frac{(c - b)^2(1 - \lambda) + (c + b)^2(1 + \lambda)}{2(1 - \lambda^2)} = \frac{b^2 + c^2 + 2\lambda bc}{1 - \lambda^2}. \end{aligned}$$

The assertion follows. $\square$

4.24 Lemma. *Let $X \in \mathbf{R}^{2\times 2}$. Then the set of all $A \in \mathbf{R}^{2\times 2}$ so that*

$$|A+tX|^2 \neq |\det(A+tX)| \quad \text{for all } t \in \mathbf{R}$$

is dense in $\mathbf{R}^{2\times 2}$.

Proof. Let $A \in \mathbf{R}^{2\times 2}$ and write

$$A = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}, \quad X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Suppose there is a $t \in \mathbf{R}$ such that $|A+tX|^2 = |\det(A+tX)|$. Setting

$$z = \frac{x_1 + x_4 + t(a+d)}{2} + i\frac{x_3 - x_2 + t(c-b)}{2}, \quad w = \frac{x_1 - x_4 + t(a-d)}{2} + i\frac{x_3 + x_2 + t(c+b)}{2},$$

it follows from Lemma 3.7 that

$$(|z|+|w|)^2 = |A+tX|^2 = |\det(A+tx)| = (|z|+|w|)||z|-|w||$$

so that

$$|z|+|w| = ||z|-|w||.$$

By taking squares on both sides, we note that this is equivalent to the assertion $|z||w| = 0$, meaning that either $z = 0$ or $w = 0$. We conclude that $|A+tX|^2 = |\det(A+tX)|$ holds if and only if we have

$$\begin{cases} x_1 + x_4 = -t(a+d) \\ x_3 - x_2 = -t(c-b), \end{cases} \quad \text{or} \quad \begin{cases} x_1 - x_4 = -t(a-d) \\ x_3 + x_2 = -t(c+b). \end{cases} \tag{4.18}$$

We proceed by cases. First suppose that $c \neq \pm b$. If a matrix $A = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$ satisfies (4.18) for some fixed $t \in \mathbf{R}$, then, for all $j \in \mathbf{N}$, we define

$$A_j := \begin{pmatrix} x_1 + \frac{1}{j} & x_2 \\ x_3 & x_4 \end{pmatrix}.$$

We claim that for some $J \in \mathbf{N}$ we have that $|A_j + sX|^2 \neq |\det(A_j + sX)|$ for all $s \in \mathbf{R}$ whenever $j > J$. Indeed, if for all $j \in \mathbf{N}$ we have that $|A_j + sX|^2 \neq |\det(A_j + sX)|$ for all $s \in \mathbf{R}$, then we can set $J = 1$. On the other hand, if there is a $J \in \mathbf{N}$ and an $s \in \mathbf{R}$ so that $|A_J + sX|^2 = |\det(A_J + sX)|$, then either $x_1 + 1/J + x_4 = -s(a+d)$ and $x_3 - x_2 = -s(c-b)$ or $x_1 + 1/J - x_4 = -s(a-d)$ and $x_3 + x_2 = -s(c+b)$. Suppose we are in the first case and suppose A satisfies the first case in (4.18). Then $-s(c-b) = x_3 - x_2 = -t(c-b)$ so that $s = t$, since we assumed that $c \neq b$. But then $x_1 + 1/J + x_4 = -t(a+d) = x_1 + x_4$, which implies that $1/J = 0$. This contradiction implies that A must satisfy the second case in (4.18). Then we have

$$2x_3 = x_3 + x_2 + x_3 - x_2 = -t(c+b) - s(c-b). \tag{4.19}$$

Now suppose that there is some $j \in \mathbf{N}$ so that there is an $s' \in \mathbf{R}$ so that $|A_j + s'X|^2 = |\det(A_j + s'X)|$. If $x_3 + x_2 = -s'(c+b)$ and $x_1 + 1/j - x_4 = -s'(a-d)$, then, as before, we may conclude that $t = s'$ and $1/j = 0$, which is absurd. Hence, we must have $x_3 - x_2 = -s'(c-b)$ and $x_1 + 1/j + x_4 = -s'(a+d)$. But then, by (4.19), we have

$$-t(c+b) - s'(c-b) = 2x_3 = x_3 + x_2 + x_3 - x_2 = -t(c+b) - s(c-b)$$

so that $s' = s$. This implies that $x_1 + 1/j + x_4 = -s(a+d) = x_1 + 1/J - x_4$ so that $j = J$. We conclude that for all $j > J$ we have $|A_j + sX|^2 \neq |\det(A_j + sX)|$ for all $s \in \mathbf{R}$. The other case is proven analogously using $c \neq -b$. This proves the claim. Since $A_{j+J} \to A$ in $\mathbf{R}^{2\times 2}$ as $j \to \infty$, this proves the assertion for the case $c \neq \pm b$.

Now assume that $c = b$. If a matrix $A = \left(\begin{smallmatrix} x_1 & x_2 \\ x_3 & x_4 \end{smallmatrix}\right)$ satisfies (4.18) for some fixed $t \in \mathbf{R}$, then, for all $j \in \mathbf{N}$, we define

$$A_j := \begin{pmatrix} x_1 & x_2 \\ x_3 + \frac{1}{j} & x_4 \end{pmatrix}.$$

As before, we claim that for some $J \in \mathbf{N}$ we have that $|A_j + sX|^2 \neq |\det(A_j + sX)|$ for all $s \in \mathbf{R}$ whenever $j > J$. Again, assume that there is a $J \in \mathbf{N}$ and an $s \in \mathbf{R}$ so that $|A_J + sX|^2 = |\det(A_J + sX)|$, then either $x_1 + x_4 = -s(a+d)$ and $x_3 + 1/J - x_2 = 0$ or $x_1 - x_4 = -s(a-d)$ and $x_3 + 1/J + x_2 = -2sb$. Suppose we are in the first case. If A satisfies the first case of (4.18), then we find that $x_3 + 1/J - x_2 = 0 = x_3 - x_2$ so that $1/J = 0$, which is absurd. We conclude that A must satisfy the second case of (4.18). Proceeding in the same way as in the case $c \neq \pm b$ we can show that if $j \in \mathbf{N}$ satisfies $|A_j + s'X|^2 \neq |\det(A_j + s'X)|$ for some $s' \in \mathbf{R}$, then we must have $s' = s$ and $j = J$. We conclude that for all $j > J$ we have $|A_j + sX|^2 \neq |\det(A_j + sX)|$ for all $s \in \mathbf{R}$, proving the claim in this case. The other case is analogous. Since $A_{j+J} \to A$ in $\mathbf{R}^{2\times 2}$ as $j \to \infty$, this proves the assertion for the case $c = b$. The case $c = -b$ is treated analogously. The assertion follows. □

We refer to [Iw2, Proposition 3.1] for a more direct proof of a generalization of this lemma to higher dimensions.

Proof of Theorem 4.19. Let $A, X \in \mathbf{R}^{2\times 2}$ with $\det X \leq 0$.

We claim that it suffices to consider the case $\gamma = |1 - p/2|$. Indeed, write $f := f_{|1-p/2|}$ and suppose $\gamma > |1 - p/2|$. It follows from Lemma 4.21 that

$$f_\gamma(t) - f(t) = \left(\gamma - \left|1 - \frac{2}{p}\right|\right) |A + tX|^p$$

is convex in $\mathbf{R}$ as a function of t. Assuming we have shown that f is convex, we may then conclude that $f_\gamma = (f_\gamma - f) + f$ is convex as the sum of two convex functions. Hence, it suffices to consider f.

Next, we claim that we may assume that $|A + tX|^2 \neq |\det(A + tX)|$ for all $t \in \mathbf{R}$. Indeed, assume we have shown the desired convexity result for such matrices and suppose we have a matrix $A \in \mathbf{R}^{2\times 2}$ so that there is some $s \in \mathbf{R}$ so that $|A + sX|^2 \neq |\det(A + sX)|$. By Lemma 4.24, we can find a sequence $(A_j)_{j\in\mathbf{N}}$ in $\mathbf{R}^{2\times 2}$ so that for all $j \in \mathbf{N}$ we have $|A_j + tX|^2 \neq |\det(A_j + tX)|$ for all $t \in \mathbf{R}$, and $A_j \to A$ in $\mathbf{R}^{2\times 2}$ as $j \to \infty$. Writing $f_j(t) := |1 - 2/p||A_j + tX|^p - |A_j + tX|^{p-2} \det(A_j + tX)$, we note that $f_j(t) \to f(t)$ for all $t \in \mathbf{R}$. Moreover, for all $x, y \in \mathbf{R}$ and $t \in [0, 1]$, our assumption implies that

$$f_j(tx + (1-t)y) \leq t f_j(x) + (1-t) f_j(y).$$

Letting $j \to \infty$ leads to convexity of f, as desired. Thus, from now on we assume that $|A+tX|^2 \neq |\det(A+tX)|$ for all $t \in \mathbf{R}$.

Next we will show that it suffices to consider the case where $|A| = 1$. First suppose $A = 0$. Then $f(t) = c|t|^p$ for $c = |1 - 2/p| - |X| \det X \geq |1 - 2/p| \geq 0$, which is convex in $\mathbf{R}$ as it is the composition of the increasing convex function $s \mapsto s^p$ in $[0, \infty[$ and the convex function $s \mapsto |s|$ in $\mathbf{R}$. Assuming $A \neq 0$, we can write

$$f(t) = |A|^{-p} \left(\left|1 - \frac{2}{p}\right| ||A|^{-1}A + t|A|^{-1}X|^p - ||A|^{-1}A + t|A|^{-1}X|^{p-2} \det(|A|^{-1}A + t|A|^{-1}X) \right).$$

Since $|A|^{-1}X$ is again a matrix of non-negative determinant, since $|A|^{-p}$ is a positive constant, and since

$$||A|^{-1}A + t|A|^{-1}X|^2 = |A|^{-2}|A+tX|^2 \neq |A|^{-2}|\det(A+tX)| = |\det(|A|^{-1}A + t|A|^{-1}X)|$$

for all $t \in \mathbf{R}$, we may replace A by $|A|^{-1}A$ and replace X by $|A|^{-1}X$ to reduce to the case $|A| = 1$.

Reducing further, we claim that we may assume that A is of the form

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}, \quad |\lambda| < 1. \tag{4.20}$$

Since, in particular, we assumed that $1 = |A|^2 = |A + 0X|^2 \neq |\det(A + 0X)| = |\det A|$, we may appeal to Lemma 4.22 to find rotation matrices $O_1, O_2 \in \mathbf{R}^{2\times 2}$ so that $O_1 A O_2$ is of the desired form. Since multiplication by rotation matrices leave operator norms and determinants invariant, noting that $O_1(A+tX)O_2 = O_1AO_2 + tO_1XO_2$, we may replace A by O_1AO_2 and X by O_1XO_2 to reduce to the case where A satisfies (4.20).

It follows from Lemma 4.23 that f is a smooth function and thus, by Lemma 4.16, to show that f is convex, we have to show that $f''(t) \geq 0$ for all $t \in \mathbf{R}$. Setting $h(t) := |A + tX|$ and

$$g(t) := \det(A+tX) = (\lambda + ta)(1+td) - t^2 bc = \lambda + (a + \lambda d)t + (\det X)t^2,$$

then g is smooth and, by Lemma 4.23, h is also smooth. By picking $t_0 \in \mathbf{R}$ and by replacing A with $A + t_0 X$ and by considering $f(t + t_0)$ instead of $f(t)$, we note that we have reduced to the case where we need only check that $f''(0) \geq 0$ to conclude that f is convex.

By the computations of the derivatives of h in Lemma 4.23 and the fact that $h(0) = |A| = 1$ we have

$$\left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right|_{t=0} (h(t))^p = p(p-1)(h(0))^{p-2}h'(0)^2 + p(h(0))^{p-1}h''(0) = p(p-1)d^2 + ph''(0).$$

Moreover, computing $g(0) = \det A = \lambda$, $g'(0) = a + \lambda d$, $g''(0) = 2\det X \leq 0$, we find that

$$\begin{aligned} &\left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right|_{t=0} (h(t))^{p-2} g(t) \\ = \;& g(0) \left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right|_{t=0} (h(t))^{p-2} + 2g'(0) \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (h(t))^{p-2} + (h(0))^{p-2} g''(0) \\ = \;& \lambda(p-2)(p-3)d^2 + \lambda(p-2)h''(0) + 2(p-2)(a+\lambda d)d + 2\det X \\ \leq \;& \lambda(p-2)(p-3)d^2 + \lambda(p-2)h''(0) + 2(p-2)(a+\lambda d)d. \end{aligned}$$

Thus,

$$f''(0) \geq \left|1 - \frac{2}{p}\right| (p(p-1)d^2 + ph''(0)) - \lambda(p-2)(p-3)d^2 - (p-2)h''(0) - 2(p-2)(a+\lambda d)d$$
$$= (|p-2|(p-1) - \lambda(p-2)(p-3))d^2 + (|p-2| - \lambda(p-2))h''(0) - 2(p-2)ad - 2\lambda(p-2)d^2.$$

Collecting the terms with d^2, we note that

$$(|p-2|(p-1) - \lambda(p-2)(p-3) - 2\lambda(p-2))d^2 = (p-1)(|p-2| - \lambda(p-2))d^2 \geq 0,$$

since $|\lambda| < 1$. Thus, by continuing our estimate, we obtain

$$f''(0) \geq (|p-2| - \lambda(p-2))h''(0) - 2(p-2)ad \geq (|p-2| - \lambda(p-2))h''(0) - 2(p-2)bc, \quad (4.21)$$

since $ad - bc = \det X \leq 0$.

We will conclude the proof by considering the cases $p \leq 2$ and $p \geq 2$ separately. Suppose $p \leq 2$. Then (4.21) becomes

$$f''(0) \geq (2-p)((1+\lambda)h''(0) + 2bc) = (2-p)\left(\frac{b^2 + c^2 + 2\lambda bc}{1-\lambda} + 2bc\right) = (2-p)\frac{(b+c)^2}{1-\lambda} \geq 0.$$

Similarly, when $p \geq 2$ we have

$$f''(0) \geq (p-2)\frac{(b-c)^2}{1+\lambda} \geq 0.$$

This proves the desired convexity result. Finally, since any rank-one matrix has determinant 0, we conclude that $E_{p,\gamma}$ is rank-one convex. The assertion follows. □

While we know that quasiconvexity implies rank-one convexity, the question wether the converse implication holds turns out to be a much more difficult problem. Let us first settle the 1-dimensional cases of this problem.

4.25 Proposition. *Let $m = 1$ or $n = 1$ and let $F : \mathbf{R}^{m \times n} \to \mathbf{R}$ be continuous. Then the following are equivalent:*

(i) F is convex;

(ii) F is quasiconvex;

(iii) F is rank-one convex.

We will use the following lemma:

4.26 Lemma. *Let $F : \mathbf{R}^{m \times n} \to \mathbf{R}$. Then F is rank-one convex if and only if*

$$F(tA + (1-t)B) \leq tF(A) + (1-t)B$$

for all $t \in [0,1]$ and all $A, B \in \mathbf{R}^{m \times n}$ such that $\operatorname{rank}(A - B) \leq 1$.

Proof. Suppose F is rank-one convex. Let $A, B \in \mathbf{R}^{m\times n}$ such that $\operatorname{rank}(A-B) \le 1$. If $\operatorname{rank}(A-B) = 0$, then $A = B$ and the result holds trivially. Thus, we may assume that $X := A - B$ has rank one. Now let $t \in [0,1]$. Then, since F is rank-one convex, we have

$$\begin{aligned} F(tA + (1-t)B) &= F(B + tX) = F(B + (t + (1-t)0)X) \\ &\le tF(B + X) + (1-t)F(B) = tF(A) + (1-t)F(B), \end{aligned}$$

as desired.

For the converse, pick $A, X \in \mathbf{R}^{m\times n}$ with $\operatorname{rank} X = 1$ and let $x, y \in \mathbf{R}$, $t \in [0,1]$. Then, since $\operatorname{rank}(A + xX - (A + yX)) = \operatorname{rank}((x-y)X) \le 1$, we have

$$\begin{aligned} F(A + (tx + (1-t)y)X) &= F(t(A + xX) + (1-t)(A + yX)) \\ &\le tF(A + xX) + (1-t)F(A + yX). \end{aligned}$$

The assertion follows. □

Proof of Proposition 4.25. The equivalence of (i) and (ii) has been established in Proposition 4.14. In view of Lemma 4.26, the equivalence of (i) and (iii) is clear because if $m = 1$ or $n = 1$, then any $A \in \mathbf{R}^{m\times n}$ satisfies $\operatorname{rank} A \le 1$. □

The notions of quasiconvexity and rank-one convexity were originally conceived in 1952, but it wasn't until 1992 that we learned that there are continuous rank-one convex functions that are not quasiconvex. In [Šv], Šverák managed to cleverly construct examples demonstrating this result for the cases where $m \ge 3$ and $n \ge 2$.

4.27 Theorem (Šverák, 1992)**.** *Let $m \ge 3$, $n \ge 2$. Then there are continuous rank-one convex functions $F : \mathbf{R}^{n\times n} \to \mathbf{R}$ that are not quasiconvex.*

While we will not carry out the full construction of Šverák's example, we will outline the idea. We define the 1-periodic function $\psi \in C^\infty(\mathbf{R}^2; \mathbf{R}^3)$ by

$$\psi(x) := \frac{1}{2\pi}(\sin 2\pi x_1, \sin 2\pi x_2, \sin 2\pi(x_1 + x_2)).$$

Then

$$D\psi(x) = \begin{pmatrix} \cos 2\pi x_1 & 0 \\ 0 & \cos 2\pi x_2 \\ \cos 2\pi(x_1 + x_2) & \cos 2\pi(x_1 + x_2) \end{pmatrix}.$$

Thus, if we define the subspace $M \subseteq \mathbf{R}^{3\times 2}$ by

$$M := \left\{ \begin{pmatrix} a & 0 \\ 0 & b \\ c & c \end{pmatrix} \,\middle|\, a, b, c \in \mathbf{R} \right\},$$

then we note that $D\psi(x) \in L$ for all $x \in \mathbf{R}^2$. We note that any rank one matrix in M must be a constant multiple of one of the spanning matrices

$$E_1 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 := \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_3 := \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}$$

of M. Thus, if we define

$$F : L \to \mathbf{R}, \quad F\left(\begin{pmatrix} a & 0 \\ 0 & b \\ c & c \end{pmatrix}\right) := -abc,$$

then, for all $x, y \in \mathbf{R}$ and $t \in [0,1]$, setting $A := aE_1 + bE_2 + cE_3$ for $a, b, c \in \mathbf{R}$, the function F satisfies

$$\begin{aligned} F(A + (tx + (1-t)y)E_1) &= -(a + (tx + (1-t)y))bc = -(t(a+x) + (1-t)(a+y))bc \\ &= tF(A + xE_1) + (1-t)F(A + yE_1) \end{aligned}$$

and similarly for E_2, E_3 instead of E_1. Hence, F is certainly convex in the direction of any rank-one matrix in M. However, we also note that since

$$\cos 2\pi(x_1 + x_2) = (\cos 2\pi x_1)(\cos 2\pi x_2) - (\sin 2\pi x_1)(\sin 2\pi x_2)$$

and since $t \mapsto (\sin 2\pi t)(\cos 2\pi t) = (\sin 4\pi t)/2$ integrates to 0 over $[0,1]$, we find that

$$\int_{[0,1]^2} F(D\psi(x))\,\mathrm{d}x = -\int_{[0,1]^2} (\cos 2\pi x_1)^2(\cos 2\pi x_2)^2\,\mathrm{d}x < 0 = F(0)$$

so that F does not satisfy the quasiconvexity condition of characterization (v) in Proposition 4.12.

Now, Šverák managed to appropriately modify F to facilitate finding an extension $F' : \mathbf{R}^{3\times 2}$ of F to all of $\mathbf{R}^{3\times 2}$ which is still rank-one convex and which is still not quasiconvex. This settles the case for $m = 3$ and $n = 2$. For the general cases where $m \geq 3$ and $n \geq 2$, we consider the surjection $P : \mathbf{R}^{m\times n} \to \mathbf{R}^{3\times 2}$ which maps a matrix $A \in \mathbf{R}^{m\times n}$ to its upper left 3×2-matrix. By considering the function $\mathbf{R}^{m\times 3} \to \mathbf{R}$, $A \mapsto F'(P(A))$ and the 1-periodic function $\psi' \in C^\infty(\mathbf{R}^n; \mathbf{R}^m)$, $\psi'(x) = (\psi(x_1, x_2), 0, \ldots, 0)$, one obtains an example of a continuous rank-one convex function that is not quasiconvex in $\mathbf{R}^{m\times n}$.

While this settles the cases $m \geq 3$, $n \geq 2$, this example does not help us in the cases where $m = 2$, which for $n \geq 2$ is still an open problem. The case $m = n = 2$ is the case we are interested in, since this is the setting of the Burkholder functional. Since we want the Burkholder functional to be quasiconvex, we formulate a conjecture as follows:

4.28 Conjecture (Morrey's Conjecture)**.** *If a continuous map $F : \mathbf{R}^{2\times 2} \to \mathbf{R}$ is rank-one convex, then F is quasiconvex.*

In the introduction of [Mo], Morrey himself seems skeptical that quasiconvexity can be characterized this way. In the book [Mo2] he merely states that it is unknown wether this result is true or false. There appears to exist evidence both for and against the validity of the conjecture, an overview of which can be found in [Ba]. For example, one may view the existence of Šverák's example as evidence against the conjecture. However, since this example is specific for the cases where $m \geq 3$, one cannot draw any conclusions for the case $m = 2$. An argument in favor of the conjecture is the fact that we can find classes of functions where rank-one convexity implies quasiconvexity. Remarkably, using the Fourier transform one can deduce this result for quadratic functions. We will present this argument here.

For $A, B \in \mathbf{C}^{m\times n}$, $A = (a_{i,k})$, $B = (b_{i,k})$, we write

$$\langle A, B\rangle := \sum_{i=1}^{n}\sum_{k=1}^{m} a_{i,k}\overline{b_{i,k}}.$$

4.29 Proposition. *Let $F : \mathbf{R}^{m\times n} \to \mathbf{R}$ and suppose F is of the form*

$$F(A) = \sum_{i,j=1}^{n}\sum_{k,l=1}^{m} c_{i,k}^{j,l} a_{i,k}a_{j,l} = \langle LA, A\rangle,$$

where $A = (a_{i,k})$ and $L = (c_{i,k}^{j,l})$ is a linear map from $\mathbf{R}^{m\times n}$ to itself that is symmetric, i.e., $c_{i,k}^{j,l} = c_{j,l}^{i,k}$ for all $i, j \in \{1, \ldots, n\}$, $j, k \in \{1, \ldots, m\}$. Then F is quasiconvex if and only if F is rank-one convex.

Proof. By Proposition 4.18 we need only show that rank-one convexity of F implies quasiconvexity of F. Suppose F is rank-one convex and let $X \in \mathbf{R}^{m\times n}$ be a rank-one matrix. Then, since F is quadratic,

$$F(X) = F\left(\left(\frac{1}{2}2 + \frac{1}{2}0\right)X\right) \leq \frac{1}{2}F(2X) + \frac{1}{2}F(0) = 2F(X).$$

Hence, since also $F(0) = 0$, we have

$$F(Y) \geq 0 \quad \text{for any } Y \in \mathbf{R}^{m\times n} \text{ with } \operatorname{rank} Y \leq 1. \tag{4.22}$$

Now let $\phi \in C_c^\infty(\mathbf{R}^n; \mathbf{R}^m)$ where $\phi = (\phi_1, \ldots, \phi_m)$ and fix $\xi \in \mathbf{R}^n$. By taking componentwise Fourier transforms, we find that

$$\mathscr{F}(D\phi)(\xi) = (2\pi i\xi_j \mathscr{F}\phi_k(\xi))_{j,k} = 2\pi i(\mathscr{F}\phi(\xi) \otimes \xi),$$

from which we conclude that $U, V \in \mathbf{R}^{m\times n}$ are rank-one matrices when U and V are respectively the real and the imaginary parts of $\mathscr{F}(D\phi)(\xi) \in \mathbf{C}^{m\times n}$. Thus, using the fact that L is symmetric, we have

$$\begin{aligned}\langle L\mathscr{F}(D\phi)(\xi), \mathscr{F}(D\phi)(\xi)\rangle &= \langle LU, U\rangle + \langle LV, V\rangle + i(\langle LV, U\rangle - \langle LU, V\rangle)\\ &= \langle LU, U\rangle + \langle LV, V\rangle = F(U) + F(V) \geq 0\end{aligned} \tag{4.23}$$

by (4.22).

Then, since the Fourier transform is a unitary isomorphism of $L^2(\mathbf{R}^n)$, we have

$$\int_{\mathbf{R}^n} \partial_i\phi_k(x)\partial_j\phi_l(x)\,\mathrm{d}x = \int_{\mathbf{R}^n} \partial_i\phi_k(x)\overline{\partial_j\phi_l(x)}\,\mathrm{d}x = \int_{\mathbf{R}^n} \mathscr{F}(\partial_i\phi_k)(\xi)\overline{\mathscr{F}(\partial_j\phi_l)(\xi)}\,\mathrm{d}\xi$$

so that

$$\int_{\mathbf{R}^n} F(D\phi(x))\,\mathrm{d}x = \int_{\mathbf{R}^n}\langle LD\phi(x), D\phi(x)\rangle\,\mathrm{d}x = \int_{\mathbf{R}^n}\langle L\mathscr{F}(D\phi)(\xi), \mathscr{F}(D\phi)(\xi)\rangle\,\mathrm{d}\xi \geq 0 \tag{4.24}$$

by (4.23).

Now pick $A \in \mathbf{R}^{m\times n}$. We note that by the Fundamental Theorem of Calculus we have

$$\int_{\mathbf{R}^n} \langle LD\phi(x), A\rangle \,\mathrm{d}x = \sum_{i,j=1}^{n} \sum_{k,l=1}^{m} c_{i,k}^{j,l} \left(\int_{\mathbf{R}^n} \partial_i \phi_k(x)\,\mathrm{d}x \right) a_{j,l} = 0$$

so that

$$\begin{aligned}\int_{\mathbf{R}^n} (F(A + D\phi(x)) - F(A))\,\mathrm{d}x &= \int_{\mathbf{R}^n} (\langle LD\phi(x), D\phi(x)\rangle + 2\langle LD\phi(x), A\rangle)\,\mathrm{d}x \\ &= \int_{\mathbf{R}^n} \langle LD\phi(x), D\phi(x)\rangle\,\mathrm{d}x \geq 0\end{aligned}$$

by (4.24). By characterization (ii) in Proposition 4.12 we conclude that F is quasiconvex. The assertion follows. □

An important example of a quadratic function is the determinant function for $m = n = 2$. Since

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \det(A + tX) = 2\det(X) = 0$$

for all $A, X \in \mathbf{R}^{2\times 2}$ where $\operatorname{rank} X = 1$, we conclude that both det and $-\det$ are rank-one convex. It then follows from Proposition 4.29 that both det and $-\det$ are quasiconvex. In particular, this implies quasiconvexity of the Burkholder functional in the case $p = 2$, since $B_2 = -\det$.

Functions $F : \mathbf{R}^{m\times n} \to \mathbf{R}$ where both F and $-F$ are quasiconvex are usually called *null Lagrangian*, which refers to the fact that all functions must solve the Euler-Lagrange equations of the integral functional associated to F. One can check that the determinant function in any dimension is null-Lagrangian, see [Ev, 8.1, Theorem 2].

Finally, we mention partial results with respect to quasiconvexity of the Burkholder functional. The following result was established in [AIPS]:

4.30 Theorem (Astala, Iwaniec, Prause, Saksman, 2010)**.** *Let $\Omega \subseteq \mathbf{C}$ be open and bounded and let $f \in W^{1,2}_{loc}(\Omega)$ be a K-quasiconformal mapping. If f extends continuously to $\overline{\Omega}$ with $f(\overline{\Omega}) = \overline{\Omega}$ and $f|_{\partial\Omega} = z$, then*

$$\int_\Omega B_p(Df(z))\,\mathrm{d}z \geq -\frac{2}{p}|\Omega| \quad \textit{for all } p \in \left[2, \frac{2K}{K-1}\right].$$

Setting $f := z + \phi$, in the critical case where $p = 2K/(K-1)$, the distortion inequality

$$|Df(z)|^2 \leq KJ_f(z) \quad \text{for a.e. } z \in \Omega,$$

which amounts to quasiconformality of f, see Proposition 3.6, is equivalent to the inequality $B_p(I + D\phi(z)) \leq 0$ for a.e. $z \in \Omega$. Astala, Iwaniec, Prause, and Saksman used this to deduce the folllowing corollary of Theorem 4.30:

4.31 Corollary. *Let $\Omega \subseteq \mathbf{R}^2$ be open and bounded and let $I \in \mathbf{R}^{2\times 2}$ denote the identity matrix. Then, under the assumption that $\phi \in C_c^\infty(\Omega; \mathbf{R}^2)$ satisfies $B_p(I + D\phi(x)) \leq 0$ for all $x \in \Omega$, we have*

$$\int_\Omega B_p(I + D\phi(x))\,\mathrm{d}x \geq \int_\Omega B_p(I)\,\mathrm{d}x = -\frac{2}{p}|\Omega|.$$

for all $p \in [2, \infty[$.

While the result does not establish quasiconvexity of B_p at the identity, it does amount to optimal gradient estimates for quasiconformal mappings as in Theorem 4.30, see [AIPS, Corollary 4.1].

Recall that we established in Proposition 4.3 that for $p \in]1, \infty[$ the functional $E : \mathbf{R}^{2\times 2} \to \mathbf{R}$ defined by

$$E(A) = E_{p, \frac{\|\mathscr{B}\|^p_{\mathcal{L}(L^p(\mathbf{C}))} - 1}{\|\mathscr{B}\|^p_{\mathcal{L}(L^p(\mathbf{C}))} + 1}}(A) = \frac{\|\mathscr{B}\|^p_{\mathcal{L}(L^p(\mathbf{C}))} - 1}{\|\mathscr{B}\|^p_{\mathcal{L}(L^p(\mathbf{C}))} + 1} |A|^p - |A|^{p-2} \det A$$

is quasiconvex at 0. Even if Iwaniec's conjecture is true, this result still does not yield quasiconvexity of the Burkholder functional at 0. However, there is an interesting application in that existing upper bounds of $\|\mathscr{B}\|^p_{\mathcal{L}(L^p(\mathbf{C}))}$ can be combined with this result to prove the following:

4.32 Proposition. *Let $U \subseteq \mathbf{C}$ be open and let $f \in W^{1,1}_{loc}(U)$. If f is orientation preserving and satisfies*

$$\int_U \frac{|Df(z)|^2}{\log(1 + |Df(z)|)} \, \mathrm{d}z < \infty,$$

then $J_f \in L^1_{loc}(U)$. Moreover, if u and v denote the respective real and imaginary parts of f, then we have the integration by parts formula

$$\begin{aligned} \int_U \phi(z) J_f(z) \, \mathrm{d}z &= \int_U u(z)(\partial_x v(z) \partial_y \phi(z) - \partial_y v(z) \partial_y \phi(z)) \, \mathrm{d}z \\ &= \int_U v(z)(\partial_x \phi(z) \partial_y u(z) - \partial_y \phi(z) \partial_y u(z)) \, \mathrm{d}z \end{aligned}$$

for all $\phi \in C_c^\infty(U)$.

We discussed in Remark 3.4 that if a map $f \in W^{1,1}_{loc}(U)$ is orientation preserving and a homeomorphism onto its image, then $J_f \in L^1_{loc}(U)$. However, the integration by parts formula may fail in this instance. This proposition gives a nice sufficient condition for both the local integrability of the Jacobian and the integration by parts formula. We refer to [AIM, Theorem 19.3.1] for a proof.

We will conclude this section with a brief discussion summarizing our findings. A summary of what we have shown is as follows:

Morrey's Conjecture$\Rightarrow$ B_p is quasiconvex $\Rightarrow$ B_p is quasiconvex at 0 $\Rightarrow$ Iwaniec's Conjecture.

While Morrey's Conjecture is the oldest conjecture here, it appears to be the least understood. It is not directly obvious how the notion of quasiconvexity is related to the notion rank-one convexity when looking at the definitions. Our deduction of the Legendre-Hadamard as a necessary requirement for existence of minimizers shows one connection, but a more direct connection is given by the fact that the Fourier transform of the Jacobian matrix of a function yields a rank-one matrix. As a matter of fact, this was the key observation in the proof that showed that the notions of rank-one convexity and quasiconvexity are equivalent for quadratic functions.

The rank-one convexity of the Burkholder functional B_p and its higher dimensional analogues together with partial results of quasiconvexity of B_p at the identity matrix and known quasiconvexity results at 0 of the functionals $E_{p,\gamma}$ for large enough γ all suggest that B_p itself may be quasiconvex, or at least quasiconvex at 0. Burkholder's estimates that show the close relation of B_p with the theory of quasiconformal mappings bear striking similarities with Pichorides' estimates with respect to his subharmonic function that eventually lead to finding the operator norms of the Hilbert transform. It is these ideas that lead the author of this thesis to believe that further study of the Burkholder functional will eventually lead to an affirmation of Iwaniec's Conjecture.

A Appendix: Convolution of Functions

In this appendix we wish to establish some fundamental results regarding the convolution of integrable functions. Our main results will be to prove Young's inequality for convolutions, see Theorem A.3 below, and to establish that, for any open $U \subseteq \mathbf{R}^n$, the space $C_c^\infty(U)$ is dense in $L^p(U)$ for all $p \in [1, \infty[$, see Theorem A.11 below.

A.1 Definition. Let $f, g : \mathbf{R}^n \to \mathbf{C}$ be measurable functions so that $y \mapsto f(y)g(x-y)$ is integrable for a.e. $x \in \mathbf{R}^n$. Then we define the *convolution product* $f * g : \mathbf{R}^n \to \mathbf{C}$ of f and g by

$$(f * g)(x) := \int_{\mathbf{R}^n} f(y)g(x-y)\,\mathrm{d}y.$$

$\diamondsuit$

The change of variables $y \mapsto x - y$ shows that for any $f, g \in L^0(\mathbf{R}^n)$ where $f * g$ is well-defined (as in the definition) we have $f * g = g * f$.

A.2 Proposition. *Let $f, g \in L^0(\mathbf{R}^n)$ so that $f * g$ is well-defined. Then $f * g \in L^0(\mathbf{R}^n)$.*

Proof. Since f, g are measurable, there exist sequences $(f_j)_{j\in\mathbf{N}}$, $(g_j)_{j\in\mathbf{N}}$ of simple functions so that $f_j \to f$ and $g_j \to g$ pointwise a.e. as $j \to \infty$. Then, for each $j \in \mathbf{N}$, the function $(x, y) \mapsto f_j(y)g_j(x)$ is again a simple function as a consequence of the formula

$$\chi_B(y)\chi_A(x) = \chi_{A\times B}(x, y),$$

for measurable sets $A, B \subseteq \mathbf{R}^n$, where χ_X denotes the indicator function of a measurable set X. By precomposing with the invertible linear transformation $(x, y) \mapsto (x - y, y)$, we conclude that

$$h_j : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{C}, \quad h(x, y) := f_j(y)g_j(x-y)$$

is a simple function for all $j \in \mathbf{N}$. Since $h_j(x, y) \to f(y)g(x-y)$ as $j \to \infty$ for a.e. $(x, y) \in \mathbf{R}^n \times \mathbf{R}^n$, we conclude that $(x, y) \mapsto f(y)g(x-y)$ is measurable. By integrating over y, we conclude from Fubini's Theorem that $f * g$ is measurable. This proves the desired result. $\square$

A.3 Theorem (Young's inequality for convolutions)**.** *Let $p, q, r \in [1, \infty]$ satisfy*

$$\frac{1}{p} + 1 = \frac{1}{q} + \frac{1}{r}.$$

Then for $f \in L^q(\mathbf{R}^n)$, $g \in L^r(\mathbf{R}^n)$, the convolution $f * g$ is well-defined. Moreover, $f * g \in L^p(\mathbf{R}^n)$ and

$$\|f * g\|_p \leq \|f\|_q \|g\|_r.$$

For the proof we require a lemma.

A.4 Lemma. *Let $f, g \in L^1(\mathbf{R}^n)$. Then $f * g$ is well-defined. Moreover, $f * g \in L^1(\mathbf{R}^n)$ and*

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1. \tag{A.1}$$

Proof. We have

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}|f(y)||g(x-y)|\,\mathrm{d}y\,\mathrm{d}x = \int_{\mathbf{R}^n}|f(y)|\int_{\mathbf{R}^n}|g(x-y)|\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbf{R}^n}|f(y)|\,\mathrm{d}y\int_{\mathbf{R}^n}|g(x)|\,\mathrm{d}x$$
$$= \|f\|_1\|g\|_1$$

so that $y \mapsto f(y)g(x-y)$ must be integrable for a.e. $x \in \mathbf{R}^n$. Moreover, this equality implies (A.1), as desired. □

Proof of Young's inequality for convolutions. First suppose that $r = \infty$. Then we must have $p = \infty$ and $q = 1$, in which case we have

$$\int_{\mathbf{R}^n}|f(y)||g(x-y)|\,\mathrm{d}y \le \|f\|_1\|g\|_\infty$$

for a.e. $x \in \mathbf{R}^n$. The asserted results follow. The case $q = \infty$ is treated analogously.

Now assume $r < \infty$ and $q < \infty$. We fix $x \in \mathbf{R}^n$ and define h_1, h_2, h_3 by

$$h_1(y) := |f(y)|^{\frac{q}{r'}}, \quad h_2(y) := |g(x-y)|^{\frac{r}{q'}}, \quad h_3(y) := |f(y)|^{\frac{q}{p}}|g(x-y)|^{\frac{r}{p}}.$$

Then $h_1 \in L^{r'}(\mathbf{R}^n)$, $h_2 \in L^{q'}(\mathbf{R}^n)$ and, moreover, we have $h_3 \in L^p(\mathbf{R}^n)$ by Lemma A.4. The relations on p, q, and r imply that we have

$$\frac{q}{r'} + \frac{q}{p} = 1, \quad \frac{r}{p} + \frac{r}{q'} = 1, \quad \frac{1}{r'} + \frac{1}{q'} + \frac{1}{p} = 1.$$

This implies that $h_1h_2h_r(y) = |f(y)||g(x-y)$ and, by Hölder's inequality for the product of three functions,

$$\int_{\mathbf{R}^n}|f(y)||g(x-y)|\,\mathrm{d}y \le \|h_1\|_{r'}\|h_2\|_p\|h_3\|_{q'} = \|f\|_q^{\frac{q}{r'}}\|g\|_r^{\frac{r}{q'}}(|f|^q * |g|^r)(x)^{\frac{1}{p}}, \tag{A.2}$$

which is finite for a.e. $x \in \mathbf{R}^n$ by Lemma A.4, which implies that $f * g$ is well-defined. Moreover, (A.1) and (A.2) imply that

$$\|f * g\|_p \le \|f\|_q^{\frac{q}{r'}}\|g\|_r^{\frac{r}{q'}}\||f|^q * |g|^r\|_1^{\frac{1}{p}} \le \|f\|_q^{\frac{q}{r'}}\|g\|_r^{\frac{r}{q'}}\|f\|_q^{\frac{q}{p}}\|g\|_r^{\frac{r}{p}} = \|f\|_q\|g\|_r.$$

The assertion follows. □

A.5 Corollary (Minkowski's inequality for convolutions)**.** *Let $p \in [1, \infty]$. If $f \in L^1(\mathbf{R}^n)$ and $g \in L^p(\mathbf{R}^n)$ then $f * g$ is well-defined. Moreover, we have $f * g \in L^p(\mathbf{R}^n)$ and*

$$\|f * g\|_p \le \|f\|_1\|g\|_p.$$

A.6 Definition. A family of functions $(f_\varepsilon)_{\varepsilon\in\mathbf{R}_+}$ in $L^1(\mathbf{R}^n)$ is called an *approximate identity* if it is uniformly bounded in $L^1(\mathbf{R})$ and satisfies

$$\int_{\mathbf{R}^n} f_\varepsilon(x)\,\mathrm{d}x = 1$$

for all $\varepsilon \in \mathbf{R}_+$ and

$$\lim_{\varepsilon\downarrow 0}\int_{|x|>r}|f_\varepsilon(x)|\,\mathrm{d}x = 0$$

for all $r \in \mathbf{R}_+$. ♢

A.7 Theorem. *Let* $(f_\varepsilon)_{\varepsilon\in\mathbf{R}_+}$ *be an approximate identity and* $p \in [1,\infty[$. *If* $g \in L^p(\mathbf{R}^n)$, *then*

$$\lim_{\varepsilon\downarrow 0} f_\varepsilon * g \to g \tag{A.3}$$

in $L^p(\mathbf{R}^n)$. *Moreover, for any* g *that is continuous on a compact subset* $K \subseteq \mathbf{R}^n$, *the limit* (A.3) *holds uniformly on* K. *If* $g \in C_0(\mathbf{R}^n) \subseteq L^\infty(\mathbf{R}^n)$, *then the limit* (A.3) *holds in* $L^\infty(\mathbf{R}^n)$.

For a proof, see [Gr, Theorem 1.2.19, Remark 1.2.22].

A.8 Proposition. *Let* $f \in L^1(\mathbf{R}^n)$ *with*

$$\int_{\mathbf{R}^n} f(x)\,\mathrm{d}x = 1.$$

For each $\varepsilon \in \mathbf{R}_+$ *we define* $f_\varepsilon \in L^1(\mathbf{R}^n)$ *by* $f_\varepsilon(x) := \varepsilon^{-n} f(x/\varepsilon)$. *Then* $(f_\varepsilon)_{\varepsilon\in\mathbf{R}_+}$ *is an approximate identity.*

Proof. Note that by the change of variables $x \mapsto \varepsilon x$ we have

$$\int_{\mathbf{R}^n} |f_\varepsilon(x)|\,\mathrm{d}x = \frac{1}{\varepsilon^n}\int_{\mathbf{R}^n} \left| f\left(\frac{x}{\varepsilon}\right)\right|\,\mathrm{d}x = \int_{\mathbf{R}^n} |f(x)|\,\mathrm{d}x = \|f\|_1$$

so that indeed $f_\varepsilon \in L^1(\mathbf{R}^n)$ for all $\varepsilon \in \mathbf{R}_+$ and the family $(f_\varepsilon)_{\varepsilon\in\mathbf{R}_+}$ is uniformly bounded by $\|f\|_1$ in $L^1(\mathbf{R}^n)$. A similar calculation shows that f_ε integrates to 1 for all $\varepsilon \in \mathbf{R}_+$, since f does.

Now let $r \in \mathbf{R}_+$. Then, for all $\varepsilon \in \mathbf{R}_+$ we denote by $\chi_{r/\varepsilon}$ the indicator function of the complement in $\mathbf{R}^n$ of the closed ball of radius r/ε centered at 0. Since this indicator function converges pointwise to 0 as $\varepsilon \downarrow 0$, we conclude from Lebesgue's Dominated Convergence Theorem that

$$\int_{|x|>r} |f_\varepsilon(x)|\,\mathrm{d}x = \int_{|\varepsilon x|>r} |f(x)|\,\mathrm{d}x = \int_{\mathbf{R}^n} \chi_{r/\varepsilon}(x)|f(x)|\,\mathrm{d}x \to 0 \quad \text{as} \quad \varepsilon \downarrow 0.$$

Hence, the family $(f_\varepsilon)_{\varepsilon\in\mathbf{R}_+}$ satisfies all the properties of an approximate identity. □

A.9 Definition. Let $U \subseteq \mathbf{R}^n$ be open and let $f \in L^0(U)$. We define the support $\operatorname{supp} f$ of f as the complement in U of the set of all points in U that have an open neighborhood $V \subseteq U$ so that $f(x) = 0$ for a.e. $x \in V$.

For $p \in [1,\infty]$ we denote by $L^p_c(U)$ the set of those $f \in L^p(U)$ where $\operatorname{supp} f$ is compact. ◊

Since the supports of two functions that are equal almost everywhere coincide, the notion of support is well-defined on the set of equivalence classes $L^0(U)$. It follows from the definition that $\operatorname{supp} f$ is closed for any $f \in L^0(U)$.

We wish to define an approximate identity consisting of compactly supported smooth functions. For this we recall that the function $\psi : \mathbf{R} \to \mathbf{R}$ defined by

$$\psi(t) := \begin{cases} e^{-\frac{1}{t}} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases}$$

lies in $C^\infty(\mathbf{R}^n)$. See also [DK, Lemma 2.7].

A.10 Definition. We define a function $\phi \in C_c^\infty(\mathbf{R}^n)$ by

$$\phi(x) := c\psi(1 - |x|^2) = \begin{cases} ce^{\frac{1}{|x|^2-1}} & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

where $c \in \mathbf{R}_+$ is chosen so that ϕ integrates to 1. For $\varepsilon \in \mathbf{R}_+$ we then set $\phi_\varepsilon(x) := \varepsilon^{-n}\phi(x/\varepsilon)$. The family $(\phi_\varepsilon)_{\varepsilon \in \mathbf{R}_+}$ is known as the *standard mollifier*. ♢

It follows from Proposition A.8 that the standard mollifier is an approximate identity. We note that for all $\varepsilon \in \mathbf{R}_+$ we have that ϕ_ε is non-negative and $\phi_\varepsilon(x) \neq 0$ if and only if $|x| < \varepsilon$.

The standard mollifier will be used to prove the following result:

A.11 Theorem. *Let $U \subseteq \mathbf{R}^n$ be open and let $p \in [1, \infty[$. Then $C_c^\infty(U)$ is dense in $L^p(U)$.*

Note that this result does not hold for $L^\infty(U)$, since the uniform limit of continuous functions is again a continuous function. In particular, the Stone-Weierstrass Theorem implies that the closure in $L^\infty(U)$ of $C_c^\infty(U)$ is $C_0(U)$.

For the proof, we require several lemmas. For two sets $A, B \subseteq \mathbf{R}^n$ we write $A + B$ for the set of those elements in $\mathbf{R}^n$ that can be written as $x + y$ with $x \in A$ and $y \in B$.

A.12 Lemma. *For any $f, g \in L^0(\mathbf{R}^n)$ where $f * g$ is well-defined we have* $\operatorname{supp}(f*g) \subseteq \overline{\operatorname{supp} f + \operatorname{supp} g}$.

Proof. Let $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^n$ so that $y \notin \operatorname{supp} f \cap (\{x\} - \operatorname{supp} g)$. Then either $y \notin \operatorname{supp} f$ or $x - y \notin \operatorname{supp} g$. This implies that $y \notin \operatorname{supp} h_x$, where $h_x \in L^1(\mathbf{R}^n)$ is defined by $h_x(y') := f(y')g(x - y')$. Hence, $\operatorname{supp} h_x \subseteq \operatorname{supp} f \cap (\{x\} - \operatorname{supp} g)$.

Now suppose $x \notin \operatorname{supp} f + \operatorname{supp} g$. Then $\operatorname{supp} f \cap (\{x\} - \operatorname{supp} g) = \emptyset$, meaning that

$$(f * g)(x) = \int_{\mathbf{R}^n} h_x(y)\,\mathrm{d}y = \int_{\operatorname{supp} f \cap (\{x\} - \operatorname{supp} g)} h_x(y)\,\mathrm{d}y = 0.$$

Thus, if $x \notin \overline{\operatorname{supp} f + \operatorname{supp} g}$, then there is some open neighborhood V of x so that $V \cap \overline{\operatorname{supp} f + \operatorname{supp} g} = \emptyset$ and hence so that $f * g$ vanishes on V. This implies that $x \notin \operatorname{supp}(f * g)$. By contraposition, this proves that $\operatorname{supp}(f * g) \subseteq \overline{\operatorname{supp} f + \operatorname{supp} g}$, as desired. □

A.13 Lemma. *Let $p \in [1, \infty]$. If $f \in L_c^p(\mathbf{R}^n)$ and $g \in C_c^\infty(\mathbf{R}^n)$, then $f * g \in C_c^\infty(\mathbf{R}^n)$.*

Proof. Since $L^p(\mathbf{R}^n) \subseteq \mathcal{S}'(\mathbf{R}^n)$ and $C_c^\infty(\mathbf{R}^n) \subseteq \mathcal{S}(\mathbf{R}^n)$, it is a consequence of Proposition B.41 below that $f * g \in C^\infty(\mathbf{R}^n)$. Hence, it suffices to check that $f * g$ has compact support. By Lemma A.12 we have $\operatorname{supp}(f * g) \subseteq \overline{\operatorname{supp} f + \operatorname{supp} g}$, so it suffices to check that the sum of two compact sets is again compact. Let $K, L \subseteq \mathbf{R}^n$ be compact. Since K and L are bounded, so is $K + L$. It remains to show that $K + L$ is closed to conclude that $K + L$ is compact. Let $(x_j)_{j \in \mathbf{N}}$ be a sequence in $K + L$ with limit $x \in \mathbf{R}^n$. Then we can find sequences $(k_j)_{j \in \mathbf{N}}$, $(l_j)_{j \in \mathbf{N}}$ so that $x_j = k_j + l_j$ for all $j \in \mathbf{N}$. Since K is compact, the sequence $(k_j)_{j \in \mathbf{N}}$ has a convergent subsequence $(k_{j_m})_{m \in \mathbf{N}}$ with limit $k \in K$. But then

$$\lim_{m \to \infty} l_{j_m} = \lim_{m \to \infty} x_{j_m} - k_{j_m} = x - k.$$

Since L is closed, we have $x - k \in L$. Hence, $x = k + l \in K + L$. This proves that $K + L$ is closed. The assertion follows. □

A.14 Lemma. *Let $U \subseteq \mathbf{R}^n$ be open and let $p \in [1, \infty[$. Then $L^p_c(U)$ is dense in $L^p(U)$.*

Proof. Let $f \in L^p(U)$ and choose a sequence of compact subsets $(K_j)_{j\in\mathbf{N}}$ of U so that $\bigcup_{j\in\mathbf{N}} K_j = U$. For each $j \in \mathbf{N}$ we denote by $\chi_j \in L^p_c(U)$ the indicator function of K_j. Then, by Lebesgue's Dominated Convergence Theorem, $\chi_j f \to f$ as $j \to \infty$ for any $f \in L^p(U)$. Since $\chi_j f \in L^p_c(U)$ for all $j \in \mathbf{N}$, the assertion follows. □

Proof of Theorem A.11. Let $f \in L^p_c(U)$. Since $K := \operatorname{supp} f \subseteq U$ is compact and disjoint from the closed set $F := \mathbf{R}^n \backslash U$, we have

$$\delta := d(K, F) := \inf_{x\in K,\, y\in F} |x - y| > 0.$$

Let $\varepsilon \leq \delta/2 \in \mathbf{R}_+$. We may view f as an element of $L^p(\mathbf{R}^n)$ by extending it by 0 outside of U. Then, by Lemma A.13, we have $f * \phi_\varepsilon \in C^\infty_c(\mathbf{R}^n)$, where $(\phi_\varepsilon)_{\varepsilon\in\mathbf{R}_+}$ is the standard mollifier.

Then, by Lemma A.12, we have

$$\operatorname{supp}(f * \phi_\varepsilon) \subseteq \overline{K + \operatorname{supp} \phi_\varepsilon} = K + \overline{B}_\varepsilon,$$

where $\overline{B}_\varepsilon$ denotes the closed ball of radius ε around 0. We claim that $K + \overline{B}_\varepsilon \subseteq U$. Indeed, let $x \in K$, $y \in \overline{B}_\varepsilon$, $z \in F$. Then

$$|x + y - z| \geq |x - z| - |y| \geq \delta - \varepsilon = \frac{\delta}{2} > 0.$$

As z was arbitrary, this implies that $x + y$ has a positive distance to F, meaning that $x + y \in U$. This proves the claim and thus that $\operatorname{supp}(f * g) \subseteq U$. We conclude that the restriction of $f * \phi_\varepsilon$ to U lies in $C^\infty_c(U)$.

But then it follows from Lemma A.7 that

$$\left(\int_U |f(x) - (f * \phi_\varepsilon)(x)|^p \,\mathrm{d}x\right)^{\frac{1}{p}} = \left(\int_{\mathbf{R}^n} |f(x) - (f * \phi_\varepsilon)(x)|^p \,\mathrm{d}x\right)^{\frac{1}{p}} \to 0 \quad \text{as} \quad \varepsilon \downarrow 0.$$

This implies that $f \in \overline{C^\infty_c(U)}$, where the closure is taken in $L^p(U)$. We conclude that $L^p_c(U) \subseteq \overline{C^\infty_c(U)}$. But then it follows from Lemma A.14 that

$$L^p(U) = \overline{L^p_c(U)} \subseteq \overline{C^\infty_c(U)} \subseteq L^p(U).$$

This proves that $\overline{C^\infty_c(U)} = L^p(U)$, proving the assertion. □

B Appendix: Distribution Theory and the Fourier Transform

This appendix will, by no means, be a comprehensive disambiguation of the theory of distributions. However, we will provide the definitions and results required for the main text. In particular, the first part of this appendix can be used as a supplement for Section 1. We refer the reader to [DK] and [Gru] for a more complete overview of the theory.

Throughout this appendix we let U be a non-empty open subset of $\mathbf{R}^n$. For each compact $K \subseteq U$ we define

$$C_K^\infty(U) := \{\phi \in C^\infty(U) \mid \operatorname{supp}\phi \subseteq K\},$$

where $\operatorname{supp}\phi$ denotes the complement of the largest open set on which ϕ vanishes, i.e.,

$$\operatorname{supp}\phi := \overline{\{x \in \mathbf{R}^n \mid \phi(x) \neq 0\}}.$$

This space becomes a locally convex Hausdorff space when equipped with the countable family of seminorms $(\rho_{K,k})_{k\in\mathbf{Z}_{\geq 0}}$ defined by

$$\rho_{K,k}(\phi) := \max_{|\alpha|\leq k} \sup_{x\in K} |\partial^\alpha \phi(x)|.$$

Here we are using the multi-index notation

$$\alpha = (\alpha_1, \ldots, \alpha_n) \in (\mathbf{Z}_{\geq 0})^n, \quad |\alpha| = \sum_{j=1}^n \alpha_j, \quad \partial^\alpha = \prod_{j=1}^n \partial_j^{\alpha_j} = \prod_{j=1}^n \frac{\partial^{\alpha_j}}{\partial x_j^{\alpha_j}}.$$

Since any sequence in $C_K^\infty(U)$ that is a Cauchy sequence with respect to all seminorms converges to a unique limit in $C_K^\infty(U)$, this space is actually a Fréchet space.

For every pair K, L of compact subsets of U such that $K \subseteq L$, we obtain a natural continuous injection

$$\iota_{KL} : C_K^\infty(U) \to C_L^\infty(U), \quad \iota_{KL}(\phi) := \phi.$$

Hence, the family $\{C_K^\infty(U) \mid K \subseteq U \text{ compact}\}$ forms a direct system over the directed set of compact subsets of U, partially ordered by inclusion. We may then form the direct limit

$$C_c^\infty(U) := \varinjlim C_K^\infty(U) = \bigcup_{K\subseteq U \text{ compact}} C_K^\infty(U).$$

This comes equipped with the largest topology so that all inclusion mappings

$$\iota_K : C_K^\infty(U) \hookrightarrow C_c^\infty(U)$$

are continuous. Hence, a functional $u : C_c^\infty(U) \to \mathbf{C}$ is continuous if and only if $u \circ \iota_K : C_K^\infty(U) \to \mathbf{C}$ is continuous for all compact $K \subseteq U$. Sometimes the space of compactly supported smooth functions with this topology is denoted by $\mathcal{D}(U)$.

This space is plenty rich. In particular, it contains *cutoff functions* that are equal to 1 on a given compact set.

B.1 Lemma. *Let $K \subseteq U$ be compact. Then there is a function $\chi \in C_c^\infty(U)$ satisfying $\chi(U) \subseteq [0, 1]$ and $\chi(x) = 1$ for all x in an open neighborhood of K.*

We refer to [DK, Corollary 2.16] for a proof. Inductively one can show the existence of partitions of unity.

B.2 Lemma. *Let $K \subseteq U$ be compact and let $\{U_j\}_{j=1}^J$ be a finite collection of open subsets of U so that $K \subseteq \bigcup_{j=1}^J U_j$. Then there exists a partition of unity $\{\chi_j\}_{j=1}^J$ subordinate to the cover $\{U_j\}_{j=1}^J$, i.e., functions $\{\chi_j\}_{j=1}^J \subseteq C_c^\infty(\mathbf{R}^n)$ that take values in $[0,1]$ such that $\sum_{j=1}^J \chi_j(x) = 1$ for all x in an open neighborhood of K and $\operatorname{supp}\chi_j \subseteq U_j$ for all $j \in \{1, \ldots, J\}$.*

We will freely use these constructions throughout the text.

B.3 Definition. We denote by $\mathcal{D}'(U)$ the space of all continuous linear functionals $u : C_c^\infty(U) \to \mathbf{C}$. Such a functional u is called a *distribution* in U. We give $\mathcal{D}'(U)$ the structure of a locally convex Hausdorff space by equipping it with the family of seminorms $(u \mapsto |u(\phi)|)$, indexed over all $\phi \in C_c^\infty(U)$. $\diamondsuit$

For a distribution $u \in \mathcal{D}'(U)$ we have different notational conventions when applying u to $\phi \in C_c^\infty(U)$, i.e., we will write either $u(\phi)$ or $\langle u, \phi\rangle$.

The following proposition is an immediate consequence of the definitions.

B.4 Proposition. *A linear functional $u : C_c^\infty(U) \to \mathbf{C}$ is a distribution in U if and only if for all compact $K \subseteq U$ there is a $c \in \mathbf{R}_+$ and a $k \in \mathbf{Z}_{\geq 0}$ so that*

$$|u(\phi)| \leq c\rho_{K,k}(\phi)$$

for all $\phi \in C_K^\infty(U)$.

In a way, distributions can be seen as generalized functions. We will make this assertion precise.

We say that an open set $V \subseteq U$ is relatively compact in U, if $\overline{V} \subseteq U$ and $\overline{V}$ is compact. As in Section 1, we define the space $L^1_{loc}(U)$ of locally integrable functions in U by

$$L^1_{loc}(U) := \{f \in L^0(U) \mid f|_V \in L^1(V) \text{ for all relatively compact } V \subseteq U\}.$$

As an immediate consequence of Lemma 1.4 we have the following result:

B.5 Lemma. *Let $f \in L^0(U)$. Then $f \in L^1_{loc}(U)$ if and only if $\phi f \in L^1(U)$ for all $\phi \in C_c^\infty(U)$.*

This allows us to give $L^1_{loc}(U)$ the structure of a locally convex Hausdorff space by equipping it with the seminorms

$$\|f\|_\phi := \|\phi f\|_1$$

for $\phi \in C_c^\infty(U)$. To verify that this space is Hausdorff, we would need to check that $\|f\|_\phi = 0$ for all $\phi \in C_c^\infty(U)$ implies that $f = 0$ in $L^1_{loc}(U)$. As the arguments needed to show this use common constructions, we briefly explain how this can be done. Choose a sequence of compact subsets $(K_j)_{j\in\mathbf{N}}$ of U so that $\bigcup_{j\in\mathbf{N}} K_j = U$. For each $j \in \mathbf{N}$ we pick a $\chi_j \in C_c^\infty(U)$ so that $\chi_j(x) = 1$ for all $x \in K_j$. Then, since $\|f\|_{\chi_j} = 0$, it follows that f vanishes a.e. on K_j. Since this holds for each $j \in \mathbf{N}$, it follows from countable subadditivity of the Lebesgue measure that $f = 0$ a.e., as desired.

Lemma B.5 serves as a motivation for the following result:

B.6 Proposition. *We define* $\iota : L^1_{loc}(U) \to \mathcal{D}'(U)$ *by*

$$\iota(f)(\phi) := \int_U f(x)\phi(x)\,\mathrm{d}x.$$

Then ι *is a well-defined continuous linear injective map.*

Proof. Linearity of ι and of $\iota(f)$ are clear. To see why $\iota(f)$ is a distribution in U, we pick a compact $K \subseteq U$ and let $\phi \in C^\infty_K(U)$. Then

$$|\iota(f)(\phi)| \le \int_U |f(x)\phi(x)|\,\mathrm{d}x = \int_K |f(x)\phi(x)|\,\mathrm{d}x \le \rho_{K,0}(\phi) \int_K |f(x)|\,\mathrm{d}x. \tag{B.1}$$

Thus, it follows from Proposition B.4 that $\iota(f)$ is indeed a distribution in U.

To see why ι is continuous, we note that the first inequality in (B.1) reads $|\iota(f)(\phi)| \le \|f\|_\phi$, which is valid for all $\phi \in C^\infty_c(U)$.

It remains to check that ι is injective. Pick $f \in L^1_{loc}(U)$ so that $\iota(f) = 0$. For any $\psi \in C^\infty_c(U)$ we have $\psi f \in L^1(U)$ by Lemma B.5. By extending ψf by 0 outside of the support of ψ, we may view it as an element of $L^1(\mathbf{R}^n)$. Let $(\phi_\varepsilon)_{\varepsilon \in \mathbf{R}_+}$ denote the standard mollifier, see Definition A.10. Then

$$(\psi f * \phi_\varepsilon)(x) = \int_U \psi(y) f(y) \phi_\varepsilon(x-y)\,\mathrm{d}y = \iota(f)(y \mapsto \psi(y)\phi(x-y)) = 0$$

for all $x \in \mathbf{R}^n$. By Lemma A.7, the function $\psi f * \phi_\varepsilon$ converges to ψf in $L^1(\mathbf{R}^n)$ as $\varepsilon \downarrow 0$ so that

$$\|f\|_\psi = \|\psi f\|_1 = \lim_{\varepsilon \downarrow 0} \|\psi f * \phi_\varepsilon\|_1 = 0.$$

As ψ was arbitrary, we conclude that $f = 0$. □

We will use the convention that whenever we say that $u \in \mathcal{D}'(U)$ is a function, we mean that there is some $f \in L^1_{loc}(U)$ so that $u = \iota(f)$. We will often drop the ι and simply write $f \in \mathcal{D}'(U)$ when $f \in L^1_{loc}(U)$.

We remark that the space $L^1_{loc}(U)$ is rather large. For example, by Hölder's inequality it continuously contains $L^p(U)$ for all $p \in [1, \infty]$ while it also contains $C^\infty_c(U)$ itself and other spaces of continuous functions. Even though functions in these spaces are not all differentiable in the classical sense, we will be generalizing the notion of differentiability to all distributions.

B.7 Definition. Let $u \in \mathcal{D}'(U)$ and let $j \in \{1, \ldots, n\}$. Then we define $\partial_j u \in \mathcal{D}'(U)$ by

$$\partial_j u(\phi) := -u(\partial_j \phi).$$

We call this a *distributional derivative* of u. ♢

We note that that distributional derivatives of continuously differentiable functions coincides with the classical derivatives by the partial integration formula. To see that the definition of distributional derivatives makes sense, we should check that $\partial_j u$ does indeed define a distribution in U for $u \in \mathcal{D}'(U)$ and $j \in \{1, \ldots, n\}$. Pick a compact $K \subseteq U$ and pick $c \in \mathbf{R}_+$, $k \in \mathbf{Z}_{\ge 0}$ so that

$$|u(\phi)| \le c\rho_{K,k}(\phi)$$

for all $\phi \in C_K^\infty(U)$. Then, since $\partial_j \phi \in C_K^\infty(U)$ for any $\phi \in C_K^\infty(U)$ and $\rho_{K,k}(\partial_j \phi) \leq \rho_{K,k+1}(\phi)$, we find that

$$|\partial_j u(\phi)| = |u(\partial_j \phi)| \leq c\rho_{K,k+1}(\phi)$$

for all $\phi \in C_K^\infty(U)$. We conclude that indeed $\partial_j u \in \mathcal{D}'(U)$.

B.8 Definition. Let $u \in \mathcal{D}'(U)$ then we define its *support* as the complement in U of the set of those points in U for which there exists an open neighborhood $V \subseteq U$ so that whenever $\phi \in C_c^\infty(U)$ satisfies $\operatorname{supp} \phi \subseteq V$, we have $u(\phi) = 0$. $\diamondsuit$

To determine the support of a distribution, we usually use the following characterization:

B.9 Proposition. *Let $u \in \mathcal{D}'(U)$. Then a closed set $F \subseteq U$ satisfies $\operatorname{supp} u \subseteq F$ if and only if for all $\phi \in C_c^\infty(U)$ satisfying $\operatorname{supp} \phi \cap F = \emptyset$, we have $u(\phi) = 0$.*

Proof. Suppose a closed set $F \subseteq U$ satisfies $\operatorname{supp} u \subseteq F$ and suppose $\phi \in C_c^\infty(U)$ satisfies $\operatorname{supp} \phi \cap F = \emptyset$. Since $\operatorname{supp} \phi$ is a compact set that is contained in the open set $U \backslash F$, by the definition of $\operatorname{supp} u$ we can find a finite cover $(V_j)_{j=1}^J$ of $\operatorname{supp} \phi$ of subsets of $U \backslash F$ so that whenever $\psi \in C_c^\infty(U)$ satisfies $\operatorname{supp} \psi \subseteq V_j$ for some $j \in \{1, \ldots, J\}$, we have $u(\psi) = 0$.

Pick a partition of unity $(\psi_j)_{j=1}^J$ in $C_c^\infty(U)$ subordinate to the cover $(V_j)_{j=1}^J$. Then, for each $j \in \{1, \ldots, J\}$, we have $\operatorname{supp} \psi_j \phi \subseteq V_j$. Hence,

$$u(\phi) = \sum_{j=1}^J u(\psi_j \phi) = 0,$$

as desired.

For the converse, suppose $F \subseteq U$ satisfies the property that for all $\phi \in C_c^\infty(U)$ satisfying $\operatorname{supp} \phi \cap F = \emptyset$, we have $u(\phi) = 0$. Pick $x \in U \backslash F$. Setting $V := U \backslash F$, we note that whenever $\phi \in C_c^\infty(U)$ satisfies $\operatorname{supp} \phi \subseteq V$, we have $\operatorname{supp} \phi \cap F = \emptyset$. Hence, for such ϕ we have $u(\phi) = 0$. By the definition of the support, we conclude that $x \in U \backslash \operatorname{supp} u$. Thus, by contraposition, $\operatorname{supp} u \subseteq F$, as desired. $\square$

Note in particular that this proposition implies that for $u \in \mathcal{D}'(U)$ we have $u = 0$ if and only if $\operatorname{supp} u = \emptyset$.

We should check that if $f \in L^1_{loc}(U)$, then its support as a function coincides with its support as a distribution.

B.10 Lemma. *Suppose $f \in L^1_{loc}(U)$. Then $\operatorname{supp} f = \operatorname{supp} \iota(f)$, where ι is defined as in Proposition B.6.*

Proof. If $x \in U \backslash \operatorname{supp} f$, then there is some open neighborhood $V \subseteq U$ of x so that f vanishes a.e. on V. This implies that for all $\phi \in C_c^\infty(U)$ with $\operatorname{supp} \phi \subseteq V$ we have

$$\langle f, \phi \rangle = \int_U f(x)\phi(x)\,\mathrm{d}x = \int_V f(x)\phi(x)\,\mathrm{d}x = 0$$

so that $x \in U \backslash \operatorname{supp} \iota(f)$. We conclude that $\operatorname{supp} \iota(f) \subseteq \operatorname{supp} f$.

For the converse, suppose $x \notin \operatorname{supp} \iota(f)$. Then there is some open neighborhood $V \subseteq U$ of x so that for all $\phi \in C_c^\infty(U)$ with $\operatorname{supp} \phi \subseteq V$ we have $\langle f, \phi \rangle = 0$. But this implies that $f|_V \in L^1_{loc}(V)$ is the zero distribution in $\mathcal{D}'(V)$. Since $L^1_{loc}(V)$ injects into $\mathcal{D}'(V)$, we conclude that f vanishes a.e. in V. Thus, $x \in U \setminus \operatorname{supp} f$. This proves the converse inclusion $\operatorname{supp} f \subseteq \operatorname{supp} \iota(f)$, proving that $\operatorname{supp} f = \operatorname{supp} \iota(f)$, as desired. □

The following example is used to characterize distributions supported in points.

B.11 Example. We assume that $0 \in U$. Then we can define the *Dirac delta distribution* $\delta \in \mathcal{D}'(U)$ by $\delta(\phi) := \phi(0)$. To see why this is a distribution, we note that for all compact $K \subseteq U$ we have

$$|\delta(\phi)| = |\phi(0)| \leq \rho_{K,0}(\phi)$$

for all $\phi \in C_K^\infty(U)$.

We claim that $\delta \notin L^1_{loc}(U)$. Indeed, if it were given by a function $f \in L^1_{loc}(U)$, then for all $\phi \in C_c^\infty(U)$ with $0 \notin \operatorname{supp} \phi$, we have

$$0 = \phi(0) = \delta(\phi) = \int_U f(x)\phi(x)\,\mathrm{d}x.$$

This implies that f vanishes a.e. in $U \setminus \{0\}$. But then f vanishes a.e. in U. Picking any $\phi \in C_c^\infty(U)$ with $\phi(0) = 1$ yields a contradiction, since then

$$1 = \delta(\phi) = \int_U f(x)\phi(x)\,\mathrm{d}x = 0.$$

This proves the claim.

Next, we will show that $\operatorname{supp} \delta = \{0\}$. In the proof of the above claim we have shown that $\operatorname{supp} \delta \subseteq \{0\}$. For the converse, let V be any open neighborhood of 0. Then we can pick a $\phi \in C_c^\infty(U)$ with $\operatorname{supp} \phi \subseteq V$ and $\phi(0) = 1$. Since $\delta(\phi) \neq 0$, we conclude that $0 \in \operatorname{supp} \delta$, proving the result. ♢

B.12 Theorem. *Suppose* $0 \in U$ *and* $u \in \mathcal{D}'(U)$ *satisfies* $\operatorname{supp} u \subseteq \{0\}$. *Then there is a* $k \in \mathbf{Z}_{\geq 0}$ *and there exist constants* $c_\alpha \in \mathbf{C}$ *for all multi-indices* α *with* $|\alpha| \leq k$ *so that*

$$u = \sum_{|\alpha| \leq k} c_\alpha \partial^\alpha \delta.$$

A proof may be found in [DK, Theorem 8.10] and is based on the Taylor expansion of smooth functions.

Before we proceed we record a general result for the product of smooth functions. For a multi-index α we write $\alpha! := \prod_{j=1}^n \alpha_j!$. For every pair of multi-indices α and β we can write $\beta \leq \alpha$ to mean that $\beta_j \leq \alpha_j$ for all $j \in \{1, \ldots, n\}$. Then, if $\beta \leq \alpha$, we may define the binomial coefficient

$$\binom{\alpha}{\beta} := \frac{\alpha!}{\beta!(\alpha - \beta)!}.$$

Then, for any $\phi, \psi \in C^\infty(U)$ and multi-index α, we have the *Leibniz rule for differentiation*

$$\partial^\alpha(\phi\psi) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta \phi \partial^{\alpha-\beta}\psi.$$

One can prove this using the Leibniz rule for the partial differential operators ∂_j, $j \in \{1, \dots, n\}$ and induction.

B.13 Definition. Let $u \in \mathcal{D}'(U)$ and $\psi \in C^\infty(U)$. Then we define $\psi u \in \mathcal{D}'(U)$ by

$$\psi u(\phi) := u(\psi\phi).$$

♢

Again, we should check that this definition makes sense by showing that ψu does indeed define a distribution for $u \in \mathcal{D}'(U)$ and $\psi \in C^\infty(U)$. Let $K \subseteq U$ be compact and let $\phi \in C_K^\infty(U)$. Pick $c \in \mathbf{R}_+$ and $k \in \mathbf{Z}_{\geq 0}$ so that $|u(\phi)| \leq c\rho_{K,k}(\phi)$. Then

$$|\psi u(\phi)| = |u(\psi\phi)| \leq c\rho_{K,k}(\psi\phi). \tag{B.2}$$

By the Leibniz rule for differentiation we have

$$\partial^\alpha(\psi\phi) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta \psi \partial^{\alpha-\beta}\phi$$

so that

$$\sup_{x \in K} |\partial^\alpha(\psi\phi)(x)| \leq \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \sup_{x \in K} |\partial^\beta \psi(x)| \rho_{K,k}(\phi)$$

for $|\alpha| \leq k$. Hence, there is a constant $c' \in \mathbf{R}_+$ so that

$$\rho_{K,k}(\psi\phi) \leq c'\rho_{K,k}(\phi).$$

By combining this with (B.2), we have indeed established that $\psi u \in \mathcal{D}'(U)$.

The product of a smooth function and a distribution still satisfy the Leibniz rule for differentiation.

B.14 Lemma. *Let $\phi \in C^\infty(U)$ and $u \in \mathcal{D}'(U)$. Then for each multi-index α we have*

$$\partial^\alpha(\phi u) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^\beta \phi \partial^{\alpha-\beta} u.$$

Proof. Let $j \in \{1, \dots, n\}$ and let $\psi \in C_c^\infty(U)$. Then $\partial_j(\phi\psi) = (\partial_j\phi)\psi + \phi\partial_j\psi$. Hence,

$$\partial_j(\phi u)(\psi) = -u(\phi\partial_j\psi) = u((\partial_j\phi)\psi) - u(\partial_j(\phi\psi)) = (\partial_j\phi)u(\psi) + \phi\partial_j u(\psi)$$

so that $\partial_j(\phi u) = (\partial_j\phi)u + \phi\partial_j u$. The general result now follows by induction. □

B.15 Proposition. *Suppose $u \in \mathcal{D}'(U)$ satisfies $\psi u = 0$ for some $\psi \in C^\infty(U)$. Then*

$$\operatorname{supp} u \subseteq \{x \in U \mid \psi(x) = 0\}.$$

Proof. Suppose $\psi(x) \neq 0$ for some $x \in U$. By continuity of ψ there is some neighborhood $V \subseteq U$ of x so that ψ is nowhere vanishing on V. Pick $\phi \in C_c^\infty(U)$ with $\operatorname{supp} \phi \subseteq V$. Then ϕ/ψ is a well-defined element of $C_c^\infty(U)$, if it is understood to vanish outside of V. Then

$$u(\phi) = \psi u \left(\frac{\phi}{\psi}\right) = 0.$$

We conclude that $x \notin \operatorname{supp} u$. The assertion follows. □

We will prove a variation of Theorem B.12.

B.16 Proposition. *Suppose $u \in \mathcal{D}'(\mathbf{R}^n)$ satisfies $x_j u = 0$ for all $j \in \{1, \ldots, n\}$. Then there is some constant $c \in \mathbf{C}$ so that $u = c\delta$.*

Proof. By Proposition B.15 we have

$$\operatorname{supp} u \subseteq \bigcap_{j=1}^{n} \{x \in \mathbf{R}^n \mid x_j = 0\} = \{0\}.$$

Now pick $\chi \in C_c^\infty(\mathbf{R}^n)$ with $\chi(x) = 1$ for x in a neighborhood of 0 and fix $\phi \in C_c^\infty(\mathbf{R}^n)$. Then $(1 - \chi(x))\phi(x) = 0$ in a neighborhood of 0 so that $\operatorname{supp}((1-\chi)\phi) \cap \operatorname{supp} u = \emptyset$. Thus, Proposition B.9 implies that

$$u(\phi) - u(\chi\phi) = u((1-\chi)\phi) = 0. \tag{B.3}$$

Now write

$$\phi(x) - \phi(0) = \int_0^1 \partial_t \phi(tx)\, \mathrm{d}t = \sum_{j=1}^{n} x_j \int_0^1 \partial_j \phi(tx)\, \mathrm{d}t$$

so that

$$\phi(x) = \phi(0) + \sum_{j=1}^{n} x_j \phi_j(x)$$

for certain $\phi_j \in C^\infty(\mathbf{R}^n)$. But then, by (B.3), we have

$$u(\phi) = u(\chi\phi) = \phi(0)u(\chi) + \sum_{j=1}^{n} x_j u(\chi\phi_j) = u(\chi)\delta(\phi).$$

This proves the assertion with $c = u(\chi)$. □

If we assume that U is invariant under dilations, i.e., $tU \subseteq U$ for all $t \in \mathbf{R}_+$, then, for any $f \in L^0(U)$ and any $t \in \mathbf{R}_+$, we can define the dilated function $d_t f \in L^0(U)$ by $d_t f(x) := f(tx)$. By transposition, we can define dilated distributions.

B.17 Definition. Suppose that U is invariant under dilations. For $u \in \mathcal{D}'(U)$ we define $d_t u \in \mathcal{D}'(U)$ by

$$d_t u(\phi) := t^{-n} u(d_{t^{-1}}\phi).$$

Moreover, we say that $u \in \mathcal{D}'(U)$ is *homogeneous of degree* $a \in \mathbf{C}$ if

$$d_t u = t^a u$$

for all $t \in \mathbf{R}_+$. $\diamondsuit$

It is straightforward to check that $d_t u$ is again a distribution for $u \in D'(U)$ and $t \in \mathbf{R}_+$ and that, for $f \in L^1_{loc}(U)$, both d_t applied to f as a function and d_t applied to f as a distribution yield the same function so that there is no ambiguity in the notation. For such f, being homogeneous at $a \in \mathbf{C}$ just means that for all $t \in \mathbf{R}_+$ we have

$$f(tx) = t^a f(x)$$

for a.e. $x \in U$. Note that typical examples of open $U \subseteq \mathbf{R}^n$ that are invariant under dilations are $\mathbf{R}^n \backslash \{0\}$ and $\mathbf{R}^n$ itself.

B.18 Example. The Dirac delta distribution $\delta \in \mathcal{D}'(\mathbf{R}^n)$ is homogeneous of degree $-n$. This is actually a special case of the following result. Let α be a multi-index and let $t \in \mathbf{R}_+$. Then

$$d_t(\partial^\alpha \delta)(\phi) = t^{-n}(-1)^{|\alpha|}\partial^\alpha(x \mapsto \phi(t^{-1}x))(0) = t^{-n-|\alpha|}(-1)^{|\alpha|}\partial^\alpha\phi(0) = t^{-n-|\alpha|}\partial^\alpha\delta(\phi).$$

Hence, $\partial^\alpha\delta$ is homogeneous of degree $-n-|\alpha|$.

As another example we consider the function $x \mapsto |x|^{-s}$ in $\mathbf{R}^n$ for $s \in \mathbf{R}$. Using spherical coordinates one can check that this function is locally integrable in $\mathbf{R}^n$ precisely when $s < n$. Using a change of variables, we find that for all $t \in \mathbf{R}_+$ and all $\phi \in C_c^\infty(\mathbf{R}^n)$ we have

$$\langle d_t|x|^{-s}, \phi\rangle = t^{-n}\int_{\mathbf{R}^n}\frac{\phi(t^{-1}x)}{|x|^s}\,\mathrm{d}x = \int_{\mathbf{R}^n}\frac{\phi(x)}{|tx|^s}\,\mathrm{d}x = t^{-s}\langle |x|^{-s}, \phi\rangle$$

so that $|x|^{-s}$ is homogeneous of degree $-s$. $\diamondsuit$

As a preparation for defining the Fourier transform, we need to define a certain class of distributions known as tempered distributions. For $x \in \mathbf{R}^n$, we write $x^\alpha := \prod_{j=1}^n x_j^{\alpha_j}$.

B.19 Definition. We define the *Schwartz space* $\mathcal{S}(\mathbf{R}^n)$ by the set of those $\phi \in C^\infty(\mathbf{R}^n)$ such that for all $m \in \mathbf{Z}_{\geq 0}$ and all multi-indices α the function $x \mapsto (1+|x|)^m\partial^\alpha\phi(x)$ is bounded in $\mathbf{R}^n$. This space is given the structure of a locally convex Hausdorff space by equipping it with the family of seminorms $(\nu_{m,k})_{m,k\in\mathbf{Z}_{\geq 0}}$ defined by

$$\nu_{m,k}(\phi) := \max_{|\alpha|\leq k}\sup_{x\in\mathbf{R}^n}(1+|x|)^m|\partial^\alpha\phi(x)|.$$

The elements of the continuous dual $\mathcal{S}'(\mathbf{R}^n)$ of $\mathcal{S}(\mathbf{R}^n)$ are called *tempered distributions*. We equip $\mathcal{S}'(\mathbf{R}^n)$ with the weak-* topology. $\diamondsuit$

We note that one can check that the Schwartz space is a Fréchet space. A typical example of a Schwartz function is $x \mapsto e^{-|x|^2}$, since this function and its derivatives vanish more quickly at infinity than polynomials. However, this also means that $x \mapsto e^{|x|^2}$ cannot define a tempered distribution as pairing off these two functions would results in a non-finite integral. However, we will later show that $\mathcal{S}'(\mathbf{R}^n)$ contains the spaces $L^p(\mathbf{R}^n)$ for $p \in [1, \infty]$ as well as the so-called space of smooth functions of moderate growth.

We note that for any $m, m', k, k' \in \mathbf{Z}_{\geq 0}$ and any $\phi \in \mathcal{S}(\mathbf{R}^n)$ we find that both $\nu_{m,k}(\phi)$ and $\nu_{m',k'}(\phi)$ are majorized by $\nu_{\max(m,m'),\max(k,k')}(\phi)$. This implies the following:

B.20 Proposition. *Let $u : \mathcal{S}(\mathbf{R}^n) \to \mathbf{C}$ be a linear functional. Then the following are equivalent:*

(i) u is a tempered distribution;

(ii) for a finite collection $(\nu_j)_{j=1}^J$ in $(\nu_{m,k})_{m,k\in\mathbf{Z}_{\geq 0}}$ there are constants $(c_j)_{j=1}^J$ in $\mathbf{R}_+$ so that

$$|u(\phi)| \leq \sum_{j=1}^{J} c_j \nu_j(\phi)$$

for all $\phi \in \mathcal{S}(\mathbf{R}^n)$;

(iii) there are $m, k \in \mathbf{Z}_{\geq 0}$ and a $c \in \mathbf{R}_+$ so that

$$|u(\phi)| \leq c\nu_{m,k}(\phi)$$

for all $\phi \in \mathcal{S}(\mathbf{R}^n)$.

The following lemma shows, in particular, why the space of Schwartz functions is a natural domain for the Fourier transform, which is defined below.

B.21 Lemma. *Let β be a multi-index. The maps*

$$\phi \mapsto x^\beta \phi, \quad \phi \mapsto \partial^\beta \phi$$

are continuous linear maps from $\mathcal{S}(\mathbf{R}^n)$ to itself. More precisely, there is a constant $c \in \mathbf{R}_+$ so that

$$\nu_{m,k}(x^\beta \phi) \leq c\nu_{m+|\beta|,k}(\phi), \quad \nu_{m,k}(\partial^\beta \phi) \leq \nu_{m,k+|\beta|}(\phi)$$

for all $m, k \in \mathbf{Z}_{\geq 0}$.

Proof. Note that for any $x \in \mathbf{R}^n$ we have

$$|x^\beta| \leq |x|^{|\beta|} \leq (1 + |x|)^{|\beta|}.$$

Moreover, for any multi-index γ we have

$$|\partial^\gamma x^\beta| = \prod_{j=1}^{n} |\partial_j^{\gamma_j} x_j^{\beta_j}| \leq c' \prod_{j=1}^{n} (1 + |x|)^{\beta_j} = c(1 + |x|)^{|\beta|}$$

for some $c' \in \mathbf{R}_+$. Thus, by the Leibniz rule for differentiation, there is a $c'' \in \mathbf{R}_+$ so that

$$|\partial^\alpha(x^\beta\phi)(x)| \le c'' \sum_{\gamma\le\alpha} |\partial^\gamma x^\alpha||\partial^{\alpha-\gamma}\phi(x)| \le c''c'(1+|x|)^{|\beta|} \sum_{\gamma\le\alpha} ||\partial^{\alpha-\gamma}\phi(x)|$$

for all $x \in \mathbf{R}^n$. Hence, for all $m, k \in \mathbf{Z}_{\ge 0}$ there is some $c \in \mathbf{R}_+$ so that for all $x \in \mathbf{R}^n$ we have

$$(1+|x|)^m|\partial^\alpha(x^\beta\phi)(x)| \le c\nu_{m+|\beta|,k}(\phi)$$

whenever $|\alpha| \le k$. We conclude that

$$\nu_{m,k}(x^\beta\phi) \le c''\nu_{m+|\beta|,k}(\phi),$$

proving the assertion about $\phi \mapsto x^\beta\phi$. The assertion about $\phi \mapsto \partial^\beta\phi$ is straightforward. The result follows. □

Next, we will show how the space of tempered distributions can be seen as a subspace of the space of distributions in $\mathbf{R}^n$. We first observe that we have $C_c^\infty(\mathbf{R}^n) \subseteq \mathcal{S}(\mathbf{R}^n)$, since any smooth function that vanishes outside of a bounded set surely vanishes quicker than any polynomial at infinity.

B.22 Lemma. *The inclusion $C_c^\infty(\mathbf{R}^n) \hookrightarrow \mathcal{S}(\mathbf{R}^n)$ is continuous. Moreover, the space $C_c^\infty(\mathbf{R}^n)$ is dense in $\mathcal{S}(\mathbf{R}^n)$.*

Proof. For the first assertion, we note that for any compact $K \subseteq \mathbf{R}^n$ and any $\phi \in C_K^\infty(\mathbf{R}^n)$ we have

$$\nu_{m,k}(\phi) \le \sup_{x\in K}(1+|x|)^m\rho_{K,k}(\phi).$$

This proves continuity of the inclusion.

For the second assertion, let $\phi \in \mathcal{S}(\mathbf{R}^n)$ be arbitrary. Pick $\chi \in C_c^\infty(\mathbf{R}^n)$ so that $\chi(\mathbf{R}^n) \subseteq [0,1]$ and $\chi(x) = 1$ whenever $|x| < 1$. Then, for $t \in \mathbf{R}_+$, we define $\phi_t := (d_t\chi)\phi \in C_c^\infty(\mathbf{R}^n)$. It suffices to show that $\lim_{t\downarrow 0}\phi_t = \phi$ in $\mathcal{S}(\mathbf{R}^n)$.

For any multi-index α we have, by the Leibniz rule for differentiation and the chain rule,

$$\partial^\alpha\phi_t(x) = \sum_{\beta\le\alpha}\binom{\alpha}{\beta}t^{|\beta|}\partial^\beta\chi(tx)\partial^{\alpha-\beta}\phi(x)$$

for $x \in \mathbf{R}^n$ and $t \in \mathbf{R}_+$. Hence, there is some $c \in \mathbf{R}_+$ so that

$$|\partial^\alpha(\phi_t-\phi)(x)| \le |(\chi(tx)-1)\partial^\alpha\phi(x)| + c\sum_{0\ne\beta\le\alpha} t^{|\beta|}|\partial^{\alpha-\beta}\phi(x)|.$$

Since $\chi(tx) - 1 = 0$ for $|x| < 1/t$, we conclude that for $t < 1$ there is a $c' \in \mathbf{R}_+$ so that

$$(1+|x|)^m|\partial^\alpha(\phi_t-\phi)(x)| \le \max_{|\alpha|\le k}\sup_{|x|\ge 1/t}(1+|x|)^m|\partial^\alpha\phi(x)| + c't\nu_{m,k}(\phi) \tag{B.4}$$

for $|\alpha| \le k$ for all $m, k \in \mathbf{Z}_{\ge 0}$. We note that

$$(1+|x|)^m|\partial^\alpha\phi(x)| \le (1+|x|)^{-1}\nu_{m+1,k}(\phi) \to 0 \quad \text{as} \quad |x| \to \infty$$

for $m \in \mathbf{Z}_{\geq 0}$ and $|\alpha| \leq k \in \mathbf{Z}_{\geq 0}$. Thus, we may conclude from (B.4) that

$$\nu_{m,k}(\phi_t - \phi) \leq \max_{|\alpha| \leq k} \sup_{|x| \geq 1/t} (1 + |x|)^m |\partial^\alpha \phi(x)| + c' t \nu_{m,k}(\phi) \to 0 \quad \text{as} \quad t \downarrow 0$$

for all $m, k \in \mathbf{Z}_{\geq 0}$. The assertion follows. □

By this lemma we see that the restriction mapping

$$\mathcal{S}'(\mathbf{R}^n) \to \mathcal{D}'(\mathbf{R}^n), \quad u \mapsto u|_{C_c^\infty(\mathbf{R}^n)}$$

is an injective continuous map, where well-definedness follows from continuity of the inclusion $C_c^\infty(\mathbf{R}^n) \hookrightarrow \mathcal{S}(\mathbf{R}^n)$ and injectivity follows from density of $C_c^\infty(\mathbf{R}^n)$ in $\mathcal{S}(\mathbf{R}^n)$. We will usually simply write $\mathcal{S}'(\mathbf{R}^n) \subseteq \mathcal{D}'(\mathbf{R}^n)$, where the identification is implied to be given by restriction. Similarly we will simply call an element of $\mathcal{D}'(U)$ a tempered distribution if it is actually the restriction of an element of $\mathcal{S}'(\mathbf{R}^n)$. An example would be the Dirac delta distribution $\delta \in \mathcal{D}'(U)$. Since

$$|\delta(\phi)| = |\phi(0)| \leq \nu_{0,0}(\phi)$$

for any $\phi \in C_c^\infty(U)$, we conclude that we actually have $\delta \in \mathcal{S}'(\mathbf{R}^n)$. As a matter of fact, one can show that any compactly supported distribution is a tempered distribution. A discussion on this can be found in [DK, p. 189].

B.23 Proposition. *Let $p \in [1, \infty]$. Then $\mathcal{S}(\mathbf{R}^n) \subseteq L^p(\mathbf{R}^n)$, where the inclusion is continuous. Furthermore, the map $\iota : L^p(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^n)$ defined by*

$$\iota(f)(\phi) := \int_{\mathbf{R}^n} f(x)\phi(x)\, \mathrm{d}x$$

is a well-defined continuous injection.

By this proposition we have the continuous inclusions

$$C_c^\infty(\mathbf{R}^n) \subseteq \mathcal{S}(\mathbf{R}^n) \subseteq L^p(\mathbf{R}^n) \subseteq \mathcal{S}'(\mathbf{R}^n) \subseteq \mathcal{D}'(\mathbf{R}^n)$$

for all $p \in [1, \infty]$. For the proof we will use an auxiliary lemma.

B.24 Lemma. *Let $s \in \mathbf{R}$. We have*

$$\int_{\mathbf{R}^n} (1 + |x|)^{-s} \mathrm{d}x < \infty$$

whenever $s > n$.

Proof. Let $c \in \mathbf{R}_+$ denote the $n-1$-dimensional Euclidean surface measure of the unit sphere in $\mathbf{R}^n$. By employing spherical coordinates we obtain

$$\int_{\mathbf{R}^n} (1 + |x|)^{-s} \mathrm{d}x = c \int_0^\infty r^{n-1}(1 + r)^{-s}\, \mathrm{d}r \leq c \int_0^\infty (1 + r)^{n-1-s}\, \mathrm{d}r = \frac{c}{s - n}$$

whenever $n - s < 0$. The assertion follows. □

Proof of Proposition B.23. First suppose $p \in [1, \infty[$. If $\phi \in \mathcal{S}(\mathbf{R}^n)$, then

$$\int_{\mathbf{R}^n} |\phi(x)|^p \, \mathrm{d}x = \int_{\mathbf{R}^n} (1+|x|)^{-(n+1)p} |(1+|x|)^{n+1}\phi(x)|^p \, \mathrm{d}x \le \nu_{n+1,0}(\phi)^p \int_{\mathbf{R}^n} (1+|x|)^{-(n+1)p} \, \mathrm{d}x < \infty$$

by Lemma B.24, so that$\phi \in L^p(\mathbf{R}^n)$. Moreover, this estimate implies that

$$\|\phi\|_p \le \left(\int_{\mathbf{R}^n} (1+|x|)^{-(n+1)p} \, \mathrm{d}x \right)^{\frac{1}{p}} \nu_{n+1,0}(\phi),$$

so that the inclusion $\mathcal{S}(\mathbf{R}^n) \subseteq L^p(\mathbf{R}^n)$ is continuous. For $p = \infty$ we simply note that $\|\phi\|_\infty = \nu_{0,0}(\phi)$.

For the next assertion, we let $p \in [1, \infty]$ and let $p' \in [1, \infty]$ denote its Hölder conjugate. For any $f \in L^p(\mathbf{R}^n)$ and any $\phi \in \mathcal{S}(\mathbf{R}^n)$ we then find, by Hölder's inequality, that

$$|\iota(f)(\phi)| \le \|f\|_p \|\phi\|_{p'}.$$

By the previous result we can estimate $\|\phi\|_{p'}$ by a constant times $\nu_{m,0}(\phi)$ for an appropriate $m \in \mathbf{Z}_{\ge 0}$, so that we may conclude that $\iota(f)$ indeed lies in $\mathcal{S}'(\mathbf{R}^n)$ and that ι is continuous.

Finally, we need to check that ι is injective. Suppose $\iota(f) = 0$ for some $f \in L^p(\mathbf{R}^n)$. Then f is a locally integrable function that defines the zero distribution in $\mathbf{R}^n$, so it must vanish a.e. in $\mathbf{R}^n$ as in the proof of Proposition B.6. The assertion follows. □

B.25 Definition. We define the space $\mathcal{O}_M(\mathbf{R}^n)$ of *smooth functions of moderate growth* to consist of those $\phi \in C^\infty(\mathbf{R}^n)$ so that for each $k \in \mathbf{Z}_{\ge 0}$ there is some $c \in \mathbf{R}_+$ and an $m_0 \in \mathbf{Z}_{\ge 0}$ so that for all multi-indices α with $|\alpha| \le k$ we have

$$|\partial^\alpha \phi(x)| \le c(1+|x|)^{m_0}$$

for all $x \in \mathbf{R}^n$. We give the space the structure of a locally convex Hausdorff space by equipping it with the seminorms

$$n_{k,\psi}(\phi) := \max_{|\alpha| \le k} \sup_{x \in \mathbf{R}^n} |\psi(x) \partial^\alpha \phi(x)|$$

for $\psi \in \mathcal{S}(\mathbf{R}^n)$. ♢

For this definition, we should first check that these seminorms are well-defined. For this we shall prove the following proposition:

B.26 Proposition. *Let $\phi \in \mathcal{O}_M(\mathbf{R}^n)$ and $\psi \in \mathcal{S}(\mathbf{R}^n)$. Then $\phi\psi \in \mathcal{S}(\mathbf{R}^n)$. Moreover, the mapping $\mathcal{S}(\mathbf{R}^n) \to \mathcal{S}(\mathbf{R}^n)$ given by $\psi \mapsto \phi\psi$ is continuous.*

Proof. Let $m, k \in \mathbf{Z}_{\ge 0}$. Since $\phi \in \mathcal{O}_M(\mathbf{R}^n)$, there is some $c \in \mathbf{R}_+$ and some $m_0 \in \mathbf{Z}_{\ge 0}$ so that for all multi-indices α with $|\alpha| \le k$ we have

$$|\partial^\alpha \phi(x)| \le c(1+|x|)^{m_0}$$

for all $x \in \mathbf{R}^n$. Hence, by the Leibniz rule of differentiation, there are $c', c'' \in \mathbf{R}_+$ so that

$$\begin{aligned}(1+|x|)^m|\partial^\alpha(\phi\psi)(x)| &\leq (1+|x|)^m \sum_{\beta\leq\alpha}\binom{\alpha}{\beta}|\partial^\beta\phi(x)||\partial^{\alpha-\beta}\psi(x)| \\ &\leq c'(1+|x|)^m\sum_{\beta\leq\alpha}c(1+|x|)^{m_0}|\partial^{\alpha-\beta}\psi(x)| \\ &\leq c''\nu_{m+m_0,k}(\psi)\end{aligned}$$

for $|\alpha| \leq k$, so that

$$\nu_{m,k}(\phi\psi) \leq c''\nu_{m+m_0,k}(\psi).$$

This proves both assertions. □

Combining this proposition with the fact that $\partial^\alpha\phi \in \mathcal{O}_M(\mathbf{R}^n)$ for any $\phi \in \mathcal{O}_M(\mathbf{R}^n)$ and any multi-index α, shows that $\psi\partial^\alpha\phi \in \mathcal{S}(\mathbf{R}^n) \subseteq L^\infty(\mathbf{R}^n)$. This proves that the seminorms on $\mathcal{O}_M(\mathbf{R}^n)$ are indeed well-defined.

To see that $\mathcal{O}_M(\mathbf{R}^n)$ is indeed Hausdorff, we should check that if for some $\phi \in \mathcal{O}_M(\mathbf{R}^n)$ we have $n_{k,\psi}(\phi) = 0$ for all $\psi \in \mathcal{S}(\mathbf{R}^n)$, then $\phi = 0$. But for this we could pick any $x \in \mathbf{R}^n$ and a $\chi \in C_c^\infty(\mathbf{R}^n)$ with $\chi(x) = 1$ so that

$$|\phi(x)| = |\phi(x)\chi(x)| \leq n_{0,\psi}(\phi) = 0,$$

showing that, since x was arbitrary, ϕ must be 0.

Next, we will now show that $\mathcal{O}_M(\mathbf{R}^n) \subseteq \mathcal{S}'(\mathbf{R}^n)$, where the inclusion is continuous.

B.27 Proposition. *We have $\mathcal{S}(\mathbf{R}^n) \subseteq \mathcal{O}_M(\mathbf{R}^n)$, where the inclusion is continuous. Furthermore, the map $\iota : \mathcal{O}_M(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^n)$ defined by*

$$\iota(\phi)(\psi) := \int_{\mathbf{R}^n}\phi(x)\psi(x)\,\mathrm{d}x \tag{B.5}$$

is a well-defined continuous injection.

Proof. Let $k \in \mathbf{Z}_{\geq 0}$ and $\phi, \psi \in \mathcal{S}(\mathbf{R}^n)$. Then for any multi-index α with $|\alpha| \leq k$ and any $x \in \mathbf{R}^n$ we have

$$|\phi(x)\partial^\alpha\psi(x)| \leq \nu_{0,0}(\phi)\nu(0,k)(\psi)$$

so that $\psi \in \mathcal{O}_M(\mathbf{R}^n)$ and

$$n_{k,\phi}(\psi) \leq \nu_{0,0}(\phi)\nu(0,k)(\psi).$$

This proves the first assertion.

Next, we will check the second assertion. First we shall check that ι is well-defined. Let $\phi \in \mathcal{O}_M(\mathbf{R}^n)$ and $\psi \in \mathcal{S}(\mathbf{R}^n)$. Then $\phi\psi \in \mathcal{S}(\mathbf{R}^n) \subseteq L^1(\mathbf{R}^n)$ by Proposition B.26 and Proposition B.23 so that the integral in (B.5) is well-defined. Moreover, these propositions imply that the composition

$$\mathcal{S}(\mathbf{R}^n) \to \mathcal{S}(\mathbf{R}^n) \hookrightarrow L^1(\mathbf{R}^n), \quad \psi \mapsto \phi\psi \mapsto \phi\psi$$

is continuous. Hence, there are $c \in \mathbf{R}_+$ and $m, k \in \mathbf{Z}_{\geq 0}$ so that

$$\|\phi\psi\|_1 \leq c\nu_{m,k}(\psi).$$

Thus,

$$|\iota(\phi)(\psi)| \leq \|\phi\psi\|_1 \leq c\nu_{m,k}(\psi),$$

proving that $\iota(\phi) \in \mathcal{S}'(\mathbf{R}^n)$, as desired.

To see that ι is continuous, we define $\chi : \mathbf{R}^n \to \mathbf{C}$ by $\chi(x) := 2^{n+1}(1+|x|^2)^{n+1}\psi(x)$ for $\psi \in \mathcal{S}(\mathbf{R}^n)$. Then $\chi \in \mathcal{S}(\mathbf{R}^n)$ by Lemma B.21 and the binomial theorem. Thus, since $(1+|x|)^2 \leq 2(1+|x|)^2$ so that

$$(1+|x|)^{2(n+1)} \leq 2^{n+1}(1+|x|^2)^{n+1},$$

we have

$$|\iota(\phi)(\psi)| \leq \int_{\mathbf{R}^n} (1+|x|)^{-2(n+1)}(1+|x|)^{2(n+1)}|\psi(x)\phi(x)| \, \mathrm{d}x \leq n_{0,\chi}(\phi) \int_{\mathbf{R}^n} (1+|x|)^{-2(n+1)} \, \mathrm{d}x.$$

By Lemma B.24, this proves the result.

Finally, we should check that ι is injective. If $\iota(\phi) = 0$ for some $\phi \in \mathcal{O}_M(\mathbf{R}^n) \subseteq L^1_{loc}(\mathbf{R}^n)$, then ϕ defines the zero distribution. Thus, we must have $\phi = 0$ a.e. in $\mathbf{R}^n$ by the corresponding result for distributions. Since ϕ is continuous, we conclude that $\phi = 0$. The assertion follows. $\square$

Next, we wish to define the Fourier transform. For any $\phi \in \mathcal{S}(\mathbf{R}^n)$ we define $\mathscr{F}\phi : \mathbf{R}^n \to \mathbf{C}$ by

$$\mathscr{F}\phi(\xi) := \int_{\mathbf{R}^n} \phi(x) e^{-2\pi i \xi \cdot x} \, \mathrm{d}x,$$

where it is customary to write ξ for the coordinates on the Fourier side.

B.28 Proposition. *The mapping $\phi \mapsto \mathscr{F}\phi$ is a well defined continuous linear map from $\mathcal{S}(\mathbf{R}^n)$ to $\mathcal{S}(\mathbf{R}^n)$. For any $\phi \in \mathcal{S}(\mathbf{R}^n)$, $\xi \in \mathbf{R}^n$, and $j \in \{1, \ldots, n\}$, we have*

$$\mathscr{F}(\partial_j \phi)(\xi) = 2\pi i \xi_j \mathscr{F}\phi(\xi), \quad \partial_j(\mathscr{F}\phi)(\xi) = -2\pi i \mathscr{F}(x_j \phi)(\xi).$$

Finally, for any pair $\phi, \psi \in \mathcal{S}(\mathbf{R}^n)$ we have

$$\langle \mathscr{F}\phi, \psi \rangle = \langle \phi, \mathscr{F}\psi \rangle. \tag{B.6}$$

Proof. Since $\mathcal{S}(\mathbf{R}^n) \subseteq L^1(\mathbf{R}^n)$ and $|\phi(x)e^{-2\pi i \xi \cdot x}| = |\phi(x)|$ for all $x, \xi \in \mathbf{R}^n$, we find that the integral that defines $\mathscr{F}$ is well-defined and that it yields a continuous function by Lebesgue's Dominated Convergence Theorem. In particular, we have $\mathscr{F}\phi \in L^\infty(\mathbf{R}^n)$, where the bound is given by $\|\phi\|_1$.

For the last assertion, we note that for any $\phi, \psi \in \mathcal{S}(\mathbf{R}^n)$ we have

$$\langle \mathscr{F}\phi, \psi \rangle = \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} \phi(x)\psi(\xi) e^{-2\pi i \xi \cdot x} \, \mathrm{d}x \, \mathrm{d}\xi = \langle \phi, \mathscr{F}\psi \rangle$$

by Fubini's Theorem. Note that all these integrals are well defined as integrals of the product of a function in $L^1(\mathbf{R}^n)$ and a function in $L^\infty(\mathbf{R}^n)$.

Next, we will check that $\mathscr{F}$ maps $\mathcal{S}(\mathbf{R}^n)$ into $\mathcal{S}(\mathbf{R}^n)$. Fix $j \in \{1, \ldots, n\}$. For any $\phi \in \mathcal{S}(\mathbf{R}^n)$ we have $x_j\phi \in \mathcal{S}(\mathbf{R}^n)$ by Lemma B.21. Hence, when differentiating with respect to ξ, we find that

$$\partial_j(\phi(x)e^{-2\pi i\xi\cdot x}) = -2\pi i x_j\phi(x)e^{-2\pi i\xi\cdot x} \tag{B.7}$$

is integrable over $\mathbf{R}^n$ with respect to x. By the theorem on differentiation under the integral sign we find that $\mathscr{F}\phi$ is partially differentiable with partial derivatives

$$\partial_j(\mathscr{F}\phi) = -2\pi i\mathscr{F}(x_j\phi) \tag{B.8}$$

by (B.7). The partial derivatives of $\mathscr{F}\phi$ are again $\mathscr{F}$ applied to a Schwartz function so that, by induction, we have $\mathscr{F}f \in C^\infty(\mathbf{R}^n)$.

Since ∂_j maps $\mathcal{S}(\mathbf{R}^n)$ into itself by Lemma B.21, we find that for any $\phi, \psi \in \mathcal{S}(\mathbf{R}^n)$ we have

$$\langle \mathscr{F}(\partial_j\phi), \psi\rangle = \langle \partial_j\phi, \mathscr{F}\psi\rangle = -\langle \phi, \partial_j\mathscr{F}\psi\rangle = \langle \phi, 2\pi i\mathscr{F}(x_j\psi)\rangle = \langle 2\pi i\xi_j\mathscr{F}\phi, \psi\rangle.$$

Since $\mathcal{S}(\mathbf{R}^n)$ is dense in $L^1(\mathbf{R}^n)$ as it contains the dense set $C_c^\infty(\mathbf{R}^n)$, we conclude that

$$\mathscr{F}(\partial_j\phi) = 2\pi i\xi_j\mathscr{F}\phi. \tag{B.9}$$

By combining the expressions (B.8) and (B.9), we find that for every pair of multi-indices α, β, we have

$$\xi^\beta\partial^\alpha(\mathscr{F}\phi) = (-1)^{|\alpha|}(2\pi i)^{|\alpha|-|\beta|}\mathscr{F}(\partial^\beta(x^\alpha\phi)) \in L^\infty(\mathbf{R}^n)$$

with bound $c\|\partial^\beta(x^\alpha\phi)\|_1$ for $c = (-1)^{|\alpha|}(2\pi i)^{|\alpha|-|\beta|}$. Thus, since for each $m \in \mathbf{Z}_{\geq 0}$ we can estimate $(1 + |x|)^m$ by a polynomial in x, we conclude that we must indeed have $\mathscr{F}\phi \in \mathcal{S}(\mathbf{R}^n)$. Moreover, by Lemma B.21 and Proposition B.23 there are $m_0, k_0 \in \mathbf{Z}_{\geq 0}$ so that for each $k \in \mathbf{Z}_{\geq 0}$ there are constants $c', c'' \in \mathbf{R}_+$ so that

$$\begin{aligned}\max_{|\alpha|\leq k}\sup_{\xi\in\mathbf{R}^n}|\xi^\beta\partial^\alpha(\mathscr{F}\phi)(\xi)| &\leq c\max_{|\alpha|\leq k}\|\partial^\beta(x^\alpha\phi)\|_1 \leq c'\max_{|\alpha|\leq k}\nu_{m_0,k_0}(\partial^\beta(x^\alpha\phi)) \\ &\leq c''\max_{|\alpha|\leq k}\nu_{m_0+|\alpha|,k_0+|\beta|}(\phi) \leq c''\nu_{m_0+k,k_0+|\beta|}(\phi).\end{aligned} \tag{B.10}$$

As each $\mathcal{S}(\mathbf{R}^n)$-seminorm of $\mathscr{F}\phi$ can be estimated by a constant times terms like the one on left-hand side of (B.10), we conclude that $\mathscr{F}$ maps $\mathcal{S}(\mathbf{R}^n)$ continuously into $\mathcal{S}(\mathbf{R}^n)$. The assertion follows. □

We can now give an important example.

B.29 Example. Define $\gamma_1 \in \mathcal{S}(\mathbf{R})$ by $\gamma_1(x) := e^{-\pi x^2}$. Then γ_1 integrates to 1, since

$$\left(\int_{\mathbf{R}} e^{-\pi x^2}\,\mathrm{d}x\right)^2 = \int_{\mathbf{R}^2} e^{-\pi(x_1^2+x_2^2)}\,\mathrm{d}x = 2\pi\int_0^\infty re^{-\pi r^2}\,\mathrm{d}r = -2\pi\left[\frac{e^{-\pi r^2}}{2\pi}\right]_0^\infty = 1.$$

Since

$$\gamma_1'(x) = -2\pi x\gamma_1(x),$$

applying $\mathscr{F}$ yields

$$2\pi i\xi\mathscr{F}\gamma_1(\xi) = \mathscr{F}(\gamma_1')(\xi) = -2\pi\mathscr{F}(x\gamma_1) = \frac{1}{i}(\mathscr{F}\gamma_1)'(\xi)$$

by Proposition B.28. Hence, $\mathscr{F}\gamma_1$ satisfies the differential equation

$$(\mathscr{F}\gamma_1)'(\xi) = -2\pi\xi\mathscr{F}\gamma_1(\xi),$$

meaning that γ_1 and $\mathscr{F}\gamma_1$ satisfy the same differential equation. But then

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathscr{F}\gamma_1(t)}{\gamma_1(t)} = -2\pi t\frac{\mathscr{F}\gamma_1(t)}{\gamma_1(t)} + 2\pi t\frac{\mathscr{F}\gamma_1(t)}{\gamma_1(t)} = 0,$$

which implies that there is a constant $c \in \mathbf{C}$ so that $\mathscr{F}\gamma_1 = c\gamma_1$. To compute c, we note that

$$c = c\gamma_1(0) = \mathscr{F}\gamma_1(0) = \int_{\mathbf{R}} \gamma_1(x)\,\mathrm{d}x = 1.$$

We conclude that $\mathscr{F}\gamma_1 = \gamma_1$.

Now define $\gamma \in \mathcal{S}(\mathbf{R}^n)$ by $\gamma(x) := e^{-\pi|x|^2}$. Then

$$\mathscr{F}\gamma(\xi) = \int_{\mathbf{R}^n} e^{-\pi|x|^2}e^{-2\pi i\xi\cdot x}\,\mathrm{d}x = \prod_{j=1}^{n}\int_{\mathbf{R}}\gamma_1(x_j)e^{-2\pi i\xi_j x_j}\,\mathrm{d}x_j = \prod_{j=1}^{n}\gamma_j(\xi_j) = \gamma(\xi)$$

for all $\xi \in \mathbf{R}^n$ so that $\mathscr{F}\gamma = \gamma$. ♢

B.30 Definition. We define the *Fourier transform* $\mathscr{F} : \mathcal{S}'(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^n)$ by

$$\mathscr{F}u(\phi) := u(\mathscr{F}\phi)$$

for $u \in \mathcal{S}'(\mathbf{R}^n)$ and $\phi \in \mathcal{S}(\mathbf{R}^n)$. ♢

By Proposition B.28 we indeed have $u \circ \mathscr{F} \in \mathcal{S}'(\mathbf{R}^n)$ for any $u \in \mathcal{S}'(\mathbf{R}^n)$ as $\mathscr{F}$ is a continuous linear mapping from $\mathcal{S}(\mathbf{R}^n)$ to itself. We also note that, by (B.6), this definition coincides with the old definition for $\mathcal{S}(\mathbf{R}^n) \subseteq \mathcal{S}'(\mathbf{R}^n)$.

B.31 Proposition. *The Fourier transform gives a homeomorphism $\mathscr{F} : \mathcal{S}(\mathbf{R}^n) \to \mathcal{S}(\mathbf{R}^n)$ whose inverse is given by $\mathscr{F}^{-1} : \mathcal{S}(\mathbf{R}^n) \to \mathcal{S}(\mathbf{R}^n)$,*

$$\mathscr{F}^{-1}\phi(x) := \mathscr{F}\phi(-x) = \int_{\mathbf{R}^n}\phi(\xi)e^{2\pi i x\cdot\xi}\,\mathrm{d}\xi$$

for $\phi \in \mathcal{S}(\mathbf{R}^n)$ and $x \in \mathbf{R}^n$.

Proof. Since the constant 1 function, which we also denote by 1, is bounded, it defines a tempered distribution. We will first show that $\mathscr{F}1 = \delta$. By Proposition B.28 we have

$$x_j\mathscr{F}1 = \frac{1}{2\pi i}\mathscr{F}(\partial_j 1) = 0$$

for all $j \in \{1, \ldots, n\}$. It follows from Proposition B.16 that $\mathscr{F}1 = c\delta$ for some $c \in \mathbf{C}$. If we define γ as in Example B.29, then

$$c = c\delta(\gamma) = \mathscr{F}1(\gamma) = \langle 1, \mathscr{F}\gamma\rangle = \langle 1, \gamma\rangle = \int_{\mathbf{R}^n} \gamma(x)\,\mathrm{d}x = 1.$$

We conclude that we indeed have $\mathscr{F}1 = \delta$.

Since the Fourier transform maps $\mathcal{S}(\mathbf{R}^n)$ continuously to itself, so does $\mathscr{F}^{-1}$. Pick any $\phi \in \mathcal{S}(\mathbf{R}^n)$ and fix $x_0 \in \mathbf{R}^n$. Then, by a change of variables, one finds

$$\mathscr{F}(x \mapsto \phi(x_0 - x))(\xi) = \mathscr{F}^{-1}\phi(\xi)e^{-2\pi i\xi\cdot x_0}.$$

Thus, we obtain

$$\begin{aligned}\phi(x_0) &= \delta(x \mapsto \phi(x_0 - x)) = \mathscr{F}1(x \mapsto \phi(x_0 - x)) \\ &= \langle 1, \mathscr{F}(x \mapsto \phi(x_0 - x))\rangle = \int_{\mathbf{R}^n} \mathscr{F}\phi(-\xi)e^{-2\pi i\xi\cdot x_0}\,\mathrm{d}\xi\end{aligned}$$

so that $\mathscr{F}(\mathscr{F}^{-1}\phi) = \phi$. But then, by a change of variables, we also have

$$\mathscr{F}^{-1}(\mathscr{F}\phi)(x) = \mathscr{F}(\mathscr{F}^{-1}(x \mapsto \phi(-x)))(-x) = \phi(x)$$

for all $x \in \mathbf{R}^n$. The assertion follows. □

Note that $\mathscr{F}^{-1} : \mathcal{S}(\mathbf{R}^n) \to \mathcal{S}(\mathbf{R}^n)$ defined as in the proposition can be extended to all of $\mathcal{S}'(\mathbf{R}^n)$ in the same way we did for the Fourier transform, i,e., for $u \in \mathcal{S}'(\mathbf{R}^n)$ we define

$$\mathscr{F}^{-1}u(\phi) := u(\mathscr{F}^{-1}\phi)$$

for $\phi \in \mathcal{S}(\mathbf{R}^n)$. This way of defining a linear operator on a space to its dual space is referred to as transposition. Using the above proposition and by unwinding the definitions we see that $\mathscr{F}^{-1}$ inverts $\mathscr{F}$ on $\mathcal{S}'(\mathbf{R}^n)$.

In the following theorem, we summarize some important properties of the Fourier transform.

B.32 Theorem. *The Fourier transform is a homeomorphism $\mathscr{F} : \mathcal{S}'(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^n)$ that restricts to a homeomorphism from $\mathcal{S}(\mathbf{R}^n)$ to $\mathcal{S}(\mathbf{R}^n)$.*

We have

$$\mathscr{F}(L^1(\mathbf{R}^n)) \subseteq L^\infty(\mathbf{R}^n), \quad \mathscr{F}(L^2(\mathbf{R}^n)) \subseteq L^2(\mathbf{R}^n),$$

where the restriction of $\mathscr{F}$ to $L^1(\mathbf{R}^n)$ satisfies $\|\mathscr{F}f\|_\infty \leq \|f\|_1$ for any $f \in L^1(\mathbf{R}^n)$ and where the restrict of $\mathscr{F}$ to $L^2(\mathbf{R}^n)$ is a unitary isomorphism of $L^2(\mathbf{R}^n)$. The same assertions hold for $\mathscr{F}^{-1}$ instead of $\mathscr{F}$.

Finally, for any $u \in \mathcal{S}'(\mathbf{R}^n)$ we have

(i) $\mathscr{F}(\partial^\alpha u) = (2\pi i)^{|\alpha|}\xi^\alpha\mathscr{F}u$, $\partial^\alpha(\mathscr{F}u) = (-2\pi i)^{|\alpha|}\mathscr{F}(x^\alpha u)$ *for every multi-index α;*

(ii) $d_t(\mathscr{F}u) = t^{-n}\mathscr{F}(d_{t^{-1}}u)$, $\mathscr{F}(d_t u) = t^{-n}d_{t^{-1}}(\mathscr{F}u)$ *for every $t \in \mathbf{R}_+$;*

(iii) $\mathscr{F}A_* = A_*\mathscr{F}$ *for any orthogonal transformation* $A \in \mathbf{R}^{n\times n}$*, where* $A_*u(\phi) := u(\phi \circ A)$ *for* $\phi \in \mathcal{S}(\mathbf{R}^n)$*;*

(iv) $\mathscr{F}(f * g) = \mathscr{F}f\mathscr{F}g$ *for* $f \in L^1(\mathbf{R}^n)$ *and* $g \in L^1(\mathbf{R}^n)$ *or* $g \in L^2(\mathbf{R}^n)$*, and similarly for* $\mathscr{F}^{-1}$ *instead of* $\mathscr{F}$*.*

Proof. To see that the Fourier transform is continuous, we note that for each $u \in \mathcal{S}'(\mathbf{R}^n)$ and each $\phi \in \mathcal{S}(\mathbf{R}^n)$, since u is a tempered distribution, there is a $c \in \mathbf{R}_+$ and $m, k \in \mathbf{Z}_{\geq 0}$ so that

$$|\mathscr{F}u(\phi)| = |u(\mathscr{F}\phi)| \leq c\nu_{m,k}(\mathscr{F}\phi).$$

Thus, since $\mathscr{F}$ maps $\mathcal{S}(\mathbf{R}^n)$ continuously to $\mathcal{S}(\mathbf{R}^n)$ by Proposition B.28, we can find a constant $c' \in \mathbf{R}_+$ and some $m', k' \in \mathbf{Z}_{\geq 0}$ so that

$$|\mathscr{F}u(\phi)| \leq c\nu_{m,k}(\mathscr{F}\phi) \leq c'\nu_{m',k'}(\phi).$$

This proves continuity of $\mathscr{F}$. Showing that $\mathscr{F}^{-1}$ is continuous is completely analogous. To see that $\mathscr{F}$ and $\mathscr{F}^{-1}$ invert each other, we use the fact that they invert each other on $\mathcal{S}(\mathbf{R}^n)$ by Proposition B.31. Then it follows that for any $u \in \mathcal{S}'(\mathbf{R}^n)$ we have

$$\mathscr{F}\mathscr{F}^{-1}u = u \circ \mathscr{F} \circ \mathscr{F}^{-1} = u,$$

and similarly $\mathscr{F}^{-1}\mathscr{F}u = u$. We conclude that $\mathscr{F}$ is indeed a homeomorphism of $\mathcal{S}'(\mathbf{R}^n)$ with inverse $\mathscr{F}^{-1}$.

To see that $\mathscr{F}(L^1(\mathbf{R}^n)) \subseteq L^\infty(\mathbf{R}^n)$, we note that

$$\int_{\mathbf{R}^n} f(x)e^{-2\pi i\xi\cdot x}\,\mathrm{d}x$$

is well defined for any $f \in L^1(\mathbf{R}^n)$ and any $\xi \in \mathbf{R}^n$. An application of Fubini's Theorem shows that this function coincides with $\mathscr{F}f$, so that the Fourier transform of an element of $L^1(\mathbf{R}^n)$ is again a function. Moreover, we find that $|\mathscr{F}f(\xi)| \leq \|f\|_1$ for all $\xi \in \mathbf{R}^n$ so that $\|\mathscr{F}f\|_\infty \leq \|f\|_1$. The result follows.

For the assertion about $L^2(\mathbf{R}^n)$, we first note that for any $\phi \in \mathcal{S}(\mathbf{R}^n)$ we have

$$\overline{\mathscr{F}\phi(\xi)} = \int_{\mathbf{R}^n} \overline{\phi(x)}e^{2\pi i\xi\cdot x}\,\mathrm{d}x = \mathscr{F}^{-1}\overline{\phi}(\xi)$$

for any $\xi \in \mathbf{R}^n$. Hence, for all $\phi, \psi \in \mathcal{S}(\mathbf{R}^n)$ we have

$$\langle\mathscr{F}\phi, \overline{\mathscr{F}\psi}\rangle = \langle\mathscr{F}\phi, \mathscr{F}^{-1}\overline{\psi}\rangle = \langle\phi, \mathscr{F}\mathscr{F}^{-1}\overline{\psi}\rangle = \langle\phi, \overline{\psi}\rangle$$

by (B.6). In particular, this means that

$$\|\mathscr{F}\phi\|_2 = \langle\mathscr{F}\phi, \overline{\mathscr{F}\phi}\rangle^{\frac{1}{2}} = \|\phi\|_2$$

for any $\phi \in \mathcal{S}(\mathbf{R}^n)$. This means that $\mathscr{F}$, viewed as a mapping from $\mathcal{S}(\mathbf{R}^n)$ to $\mathcal{S}(\mathbf{R}^n)$, has an isometric extension $F : L^2(\mathbf{R}^n) \to L^2(\mathbf{R}^n)$. Since $L^2(\mathbf{R}^n) \subseteq \mathcal{S}'(\mathbf{R}^n)$, we should check that $Ff =$

$\mathscr{F}f$ for any $f \in L^2(\mathbf{R}^n)$. Fix $f \in L^2(\mathbf{R}^n)$. As $\mathcal{S}(\mathbf{R}^n)$ is dense in $L^2(\mathbf{R}^n)$, since it contains the dense set $C_c^\infty(\mathbf{R}^n)$, we can pick a sequence $(\phi_j)_{j\in\mathbf{N}}$ in $\mathcal{S}(\mathbf{R}^n)$ that converges to f in $L^2(\mathbf{R}^n)$. Since F is continuous, this means that $(F\phi_j)_{j\in\mathbf{N}}$ converges in $L^2(\mathbf{R}^n)$ to Ff. As the inclusion $L^2(\mathbf{R}^n) \subseteq \mathcal{S}'(\mathbf{R}^n)$ is continuous, we conclude that $(F\phi_j)_{j\in\mathbf{N}}$ converges to Ff in $\mathcal{S}'(\mathbf{R}^n)$. On the other hand, we can also conclude that $(\phi_j)_{j\in\mathbf{N}}$ converges to f in $\mathcal{S}'(\mathbf{R}^n)$. Since $\mathscr{F}$ is continuous as a map from $\mathcal{S}'(\mathbf{R}^n)$ to $\mathcal{S}'(\mathbf{R}^n)$, we conclude that also $(\mathscr{F}\phi_j)_{j\in\mathbf{N}}$ converges to $\mathscr{F}f$ in $\mathcal{S}'(\mathbf{R}^n)$. But since $\mathscr{F}\phi_j = F\phi_j$ for all $j \in \mathbf{N}$, we conclude, since $\mathcal{S}'(\mathbf{R}^n)$ is Hausdorff, that

$$Ff = \lim_{j\to\infty} F\phi_j = \lim_{j\to\infty} \mathscr{F}\phi_j = \mathscr{F}f$$

where the limits are in $\mathcal{S}'(\mathbf{R}^n)$. But this means that for any $f \in L^2(\mathbf{R}^n)$ we have $\mathscr{F}f = Ff \in L^2(\mathbf{R}^n)$ with

$$\|\mathscr{F}f\|_2 = \|f\|_2.$$

Showing that a similar result holds for $\mathscr{F}^{-1}$ is analogous. We conclude that $\mathscr{F}$ restricts to a unitary isomorphism of $L^2(\mathbf{R}^n)$.

The assertion (i) follows by transposition and Proposition B.31. For (ii) and (iii), one first checks their validity for functions in $\mathcal{S}(\mathbf{R}^n)$ by applying a suitable change of variables. The general results follow by transposition. We will give a proof of (iii) and omit the proof of (ii).

We first note that for any orthogonal transformation $A \in \mathbf{R}^{n\times n}$ we have that $\phi \mapsto \phi \circ A$ leaves the seminorms on $\mathcal{S}(\mathbf{R}^n)$ invariant so that indeed $A_*u \in \mathcal{S}'(\mathbf{R}^n)$ for any $u \in \mathcal{S}'(\mathbf{R}^n)$. Fix $\phi \in \mathcal{S}(\mathbf{R}^n)$. Using the change of variables $x \mapsto A^t x = A^{-1}x$ we obtain, since $|\det A| = 1$,

$$\begin{aligned}\mathscr{F}(\phi \circ A)(\xi) &= \int_{\mathbf{R}^n} \phi(Ax)e^{-2\pi i\xi\cdot Ax}\,\mathrm{d}x = \int_{\mathbf{R}^n} \phi(x)e^{-2\pi i\xi\cdot A^t x}\,\mathrm{d}x \\ &= \int_{\mathbf{R}^n} \phi(x)e^{-2\pi iA\xi\cdot x}\,\mathrm{d}x = \mathscr{F}\phi(A\xi)\end{aligned}$$

for all $\xi \in \mathbf{R}^n$ so that $\mathscr{F}(\phi \circ A) = \mathscr{F}\phi \circ A$. Thus, for any $u \in \mathcal{S}'(\mathbf{R}^n)$ we have

$$\mathscr{F}A_*u(\phi) = u(\mathscr{F}\phi \circ A) = u(\mathscr{F}(\phi \circ A)) = A_*\mathscr{F}u(\phi).$$

This proves (iii).

For (iv), we first note that for any $f \in L^1(\mathbf{R}^n)$ and $g \in L^1(\mathbf{R}^n)$ or $g \in L^2(\mathbf{R}^n)$ we have $f * g \in L^1(\mathbf{R}^n)$ or $f * g \in L^2(\mathbf{R}^n)$ respectively by Minkowski's inequality for convolutions, see Lemma A.5. First we assume that $f, g \in L^1(\mathbf{R}^n)$. Then, since

$$e^{-2\pi\xi\cdot x} = e^{-2\pi\xi\cdot x-y}e^{-2\pi\xi\cdot y}$$

for all $x, y, \xi \in \mathbf{R}^n$, we have

$$\mathscr{F}(f*g)(\xi) = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} f(x-y)e^{-2\pi\xi\cdot x-y}g(y)e^{-2\pi\xi\cdot y}\,\mathrm{d}y\,\mathrm{d}x = \mathscr{F}f(\xi)\mathscr{F}g(\xi)$$

for all $\xi \in \mathbf{R}^n$ by Fubini's Theorem.

Now suppose $f \in L^1(\mathbf{R}^n)$ and $g \in L^2(\mathbf{R}^n)$. Since $L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$ contains $C_c^\infty(\mathbf{R}^n)$, this space is dense in $L^2(\mathbf{R}^n)$. Pick a sequence $(g_j)_{j\in\mathbf{N}}$ in $L^1(\mathbf{R}^n) \cap L^2(\mathbf{R}^n)$ that converges to g in $L^2(\mathbf{R}^n)$.

By Minkowski's inequality for convolutions we have $f * g_j \to f * g$ as $j \to \infty$ in $L^2(\mathbf{R}^n)$. As $\mathscr{F}$ maps $L^2(\mathbf{R}^n)$ isometrically into itself, we conclude that $\mathscr{F} f \mathscr{F} g_j = \mathscr{F}(f * g_j)$ converges to $\mathscr{F}(f * g)$ as $j \to \infty$ in $L^2(\mathbf{R}^n)$. On the other hand, since $\mathscr{F} g_j \in L^2(\mathbf{R}^n)$ and since multiplication by an essentially bounded function is a continuous operation in $L^2(\mathbf{R}^n)$, we find that $\mathscr{F} f \mathscr{F} g_j \to \mathscr{F} f \mathscr{F} g$ in $L^2(\mathbf{R}^n)$. Taking limits in $L^2(\mathbf{R}^n)$, we conclude that

$$\mathscr{F}(f * g) = \lim_{j \to \infty} \mathscr{F}(f * g_j) = \lim_{j \to \infty} \mathscr{F} f \mathscr{F} g_j = \mathscr{F} f \mathscr{F} g.$$

This proves the desired result. The assertions for $\mathscr{F}^{-1}$ are proven analogously. $\square$

Properties (ii) and (iii) imply the following result:

B.33 Corollary. *A tempered distribution $u \in \mathcal{S}'(\mathbf{R}^n)$ is rotationally invariant, i.e. $A_* u = u$ for all orthogonal transformations $A \in \mathbf{R}^{n \times n}$ if and only if $\mathscr{F} u$ is rotationally invariant. Moreover, u is homogeneous of degree $a \in \mathbf{C}$ if and only if $\mathscr{F} u$ is homogeneous of degree $-n-a$.*

B.34 Example. Let α be a multi-index. Then we wish to compute the Fourier transform of $\partial^\alpha \delta \in \mathcal{S}'(\mathbf{R}^n)$. First note that

$$\mathscr{F}\delta(\phi) = \mathscr{F}\phi(0) = \int_{\mathbf{R}^n} \phi(x)\,\mathrm{d}x = \langle 1, \phi \rangle$$

for all $\phi \in \mathcal{S}(\mathbf{R}^n)$, so that $\mathscr{F}\delta = 1$. Therefore, we have

$$\mathscr{F}(\partial^\alpha \delta) = (2\pi i)^{|\alpha|} \xi^\alpha \mathscr{F}\delta = (2\pi i)^{|\alpha|} \xi^\alpha.$$

Moreover, we can also compute

$$\mathscr{F}^{-1}(\partial^\alpha \delta) = (-2\pi i)^{|\alpha|} x^\alpha.$$

This means that if $u \in \mathcal{S}'(\mathbf{R}^n)$ is a distribution satisfying $\operatorname{supp} \mathscr{F} u \subseteq \{0\}$, then it follows from Theorem B.12 that $\mathscr{F} u$ is a linear combination of terms of the form $\partial^\alpha \delta$. Hence, by applying $\mathscr{F}^{-1}$ we conclude that u is a polynomial.

We describe a typical scenario where this would occur. Let $P : \mathbf{R}^n \to \mathbf{C}$ be a polynomial, i.e.,

$$P(x) := \sum_{|\alpha| \le k} c_\alpha x^\alpha$$

for $k \in \mathbf{Z}_{\ge 0}$ and $c_\alpha \in \mathbf{C}$. We then write

$$D_j := \frac{1}{2\pi i} \partial_j, \quad P(D) := \sum_{|\alpha| \le k} \frac{c_\alpha}{(2\pi i)^{|\alpha|}} \partial^\alpha \tag{B.11}$$

for the associated partial differential operator. Now assume that $P(x) = 0$ if and only if $x = 0$ and suppose $u \in \mathcal{S}'(\mathbf{R}^n)$ satisfies $P(D)u = 0$. Then

$$0 = \mathscr{F}(P(D)u) = P(\xi)u,$$

so that, by Proposition B.15, we have $\operatorname{supp} \mathscr{F} u \subseteq \{0\}$. We conclude that any $u \in \mathcal{S}'(\mathbf{R}^n)$ that satisfies $P(D)u = 0$ must be a polynomial. We can, for example, take $P(D)$ to be the Laplacian Δ, where

$$\Delta := -\sum_{j=1}^{n} \partial_j^2,$$

since $\mathscr{F}\Delta = (2\pi)^2|\xi|^2$. In the case where $n = 2$, two important examples are the Wirtinger derivatives

$$\partial_z := \frac{1}{2}(\partial_1 - i\partial_2), \quad \partial_{\overline{z}} := \frac{1}{2}(\partial_1 + i\partial_2).$$

These satisfy $\mathscr{F}\partial_z = \pi i(\xi_1 + i\xi_2)$, $\mathscr{F}\partial_{\overline{z}} = \pi i(\xi_1 - i\xi_2)$ which both only vanish when $\xi = 0$. We refer to Appendix C for a more complete discussion of the Wirtinger derivatives. ♢

The Laplacian, as well as the Wirtinger derivatives, are examples of so-called elliptic linear partial differential operators. These have the particularly nice property that certain distributional solutions to associated differential equations must automatically also be classical solutions in the sense that these distributions must be given by classically differentiable functions. This is referred to as elliptic regularity. We will make these statements more precise.

B.35 Definition. For $u \in \mathcal{D}'(U)$ we define its *singular support* $\operatorname{sing\,supp} u$ as the complement in U of the set of all points in U that have an open neighborhood $V \subseteq U$ so that there is a smooth function $\psi \in C_c^\infty(V)$ so that for any $\phi \in C_c^\infty(U)$ with $\operatorname{supp} \phi \subseteq V$ we have

$$u(\phi) = \int_V \psi(x)\phi(x)\,\mathrm{d}x.$$

♢

The singular support of a distribution $u \in \mathcal{D}'(U)$ is the complement of the largest open set where u is given by a smooth function. In particular this means that $\operatorname{sing\,supp} u = \emptyset$ if and only if u is a smooth function. Since the zero function is smooth, we have $\operatorname{sing\,supp} u \subseteq \operatorname{supp} u$ for any $u \in \mathcal{D}'(U)$.

As an example, the only singular point of δ is 0 so that $\operatorname{sing\,supp} \delta = \{0\}$.

B.36 Definition. Let $P : \mathbf{R}^n \to \mathbf{C}$ be a polynomial,

$$P(x) := \sum_{|\alpha| \le k} c_\alpha x^\alpha$$

for $k \in \mathbf{Z}_{\ge 0}$, $c_\alpha \in \mathbf{C}$. We define the associated linear partial differential operator $P(D)$ as in (B.11). We say that this linear partial differential operator is *elliptic* if

$$\sum_{|\alpha| = k} c_\alpha x^\alpha = 0 \quad \text{if and only if} \quad x = 0.$$

♢

Note that ellipticity of a linear partial differential operator only depends on its highest order terms.

B.37 Theorem (Elliptic Regularity)**.** *Suppose $P(D)$ is an elliptic partial differential operator. For any $u \in \mathcal{D}'(U)$ we have*

$$\operatorname{sing\,supp} P(D)u = \operatorname{sing\,supp} u.$$

In particular, if $P(D)u = 0$, then u is a smooth function.

For a proof, see [DK, Theorem 17.6]. So much for our discussion on elliptic regularity.

B.38 Theorem. *Let $u \in \mathcal{S}'(\mathbf{R}^n)$. Suppose that u is homogeneous of any degree and that* $\operatorname{sing\,supp} u \subseteq \{0\}$. *Then* $\operatorname{sing\,supp} \mathscr{F}u \subseteq \{0\}$.

See [Gr, Proposition 2.4.8] for a proof of this result. The proof uses an appropriate splitting of the distribution and the fact that the Fourier transform of a compactly supported distribution is given by a smooth function, the proof of which is given in [Gr, Theorem 2.3.21].

Let $x \in \mathbf{R}^n$. Then we define the reflected translation $T^x : \mathbf{R}^n \to \mathbf{R}^n$ by $T^x(y) := x - y$.

B.39 Lemma. *The mapping $\phi \mapsto \phi \circ T^x$ is a continuous linear mapping from $\mathcal{S}(\mathbf{R}^n)$ to itself. Moreover, for each $u \in \mathcal{S}'(\mathbf{R}^n)$ the mapping $T^x_* u : \mathcal{S}(\mathbf{R}^n) \to \mathbf{C}$ defined by*

$$T^x_* u(\phi) := u(\phi \circ T^x)$$

is a tempered distribution.

Proof. Note that for each $x, y \in \mathbf{R}^n$ we have

$$1 + |x + y| \leq 1 + |x| + |y| \leq 1 + |x| + |y| + |x||y| = (1 + |x|)(1 + |y|).$$

Thus, for all $m, k \in \mathbf{Z}_{\geq 0}$ we have

$$\nu_{m,k}(\phi \circ T^x) = \max_{|\alpha| \leq k} \sup_{y \in \mathbf{R}^n} (1 + |x + y|)^m |\partial^\alpha \phi(y)| \leq (1 + |x|)^m \nu_{m,k}(\phi). \tag{B.12}$$

This proves the first assertion.

For the second assertion we let $u \in \mathcal{S}'(\mathbf{R}^n)$. Picking $m, k \in \mathbf{Z}_{\geq 0}$ and $c \in \mathbf{R}_+$ so that $|u(\phi)| \leq c\nu_{m,k}(\phi)$ for all $\phi \in \mathcal{S}(\mathbf{R}^n)$, we conclude from (B.12) that

$$|T^x_* u(\phi)| = |u(\phi \circ T^x)| \leq c\nu_{m,k}(\phi \circ T^x) \leq c(1 + |x|)^m \nu_{m,k}(\phi) \quad \text{for all } \phi \in \mathcal{S}(\mathbf{R}^n). \tag{B.13}$$

The assertion follows. □

B.40 Definition. For any pair $u \in \mathcal{S}'(\mathbf{R}^n)$, $\phi \in \mathcal{S}(\mathbf{R}^n)$ we define the *convolution* $u * \phi : \mathbf{R}^n \to \mathbf{C}$ of u and ϕ by

$$(u * \phi)(x) := T^x_* u(\phi),$$

where $T^x_* u$ is defined as in Lemma B.39. ♢

One can similarly define the convolution of a distribution and a compactly supported smooth function in which case one ends up with a smooth function. For the relevant properties of such convolutions we refer to [DK, Chapter 11]. As an important example, we observe that for any $\phi \in \mathcal{S}(\mathbf{R}^n)$ we have

$$(\delta * \phi)(x) = \phi(T^x(0)) = \phi(x)$$

for all $x \in \mathbf{R}^n$ so that $\delta * \phi = \phi$.

B.41 Proposition. *Let $u \in \mathcal{S}'(\mathbf{R}^n)$ and $\phi \in \mathcal{S}(\mathbf{R}^n)$. Then $u * \phi \in \mathcal{O}_M(\mathbf{R}^n)$. For any multi-index α we have*

$$\partial^\alpha(u * \phi) = \partial^\alpha u * \phi = u * \partial^\alpha \phi. \tag{B.14}$$

Moreover, the convolution operator

$$u* : \mathcal{S}(\mathbf{R}^n) \to \mathcal{O}_M(\mathbf{R}^n), \quad \psi \mapsto u * \psi$$

is a continuous linear mapping. Finally, we have

$$\mathscr{F}(u * \phi) = \mathscr{F}\phi \mathscr{F}u$$

as the product of a smooth function and a distribution.

Proof. We will first prove (B.14). We proceed in steps.

1. We will first show that $\phi \circ T^{-x} \to \phi$ as $x \to 0$ in $\mathcal{S}(\mathbf{R}^n)$. Pick $\phi \in \mathcal{S}(\mathbf{R}^n)$. For all $x, y \in \mathbf{R}^n$ we write

$$\phi(x+y) - \phi(y) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\phi(sx+y)\,\mathrm{d}s = \sum_{j=1}^n x_j \int_0^1 \partial_j\phi(sx+y)\,\mathrm{d}s$$

so that

$$|(\phi \circ T^{-x})(y) - \phi(y)| \le \sum_{j=1}^n |x_j| \int_0^1 |\partial_j \phi(sx+y)|\,\mathrm{d}s \le |x| \sum_{j=1}^n \int_0^1 |(\partial_j\phi \circ T^{-sx})(y)|\,\mathrm{d}s. \tag{B.15}$$

Since, for every multi-index α and $m \in \mathbf{Z}_{\ge 0}$, we have

$$\begin{aligned}(1+|y|)^m \int_0^1 |(\partial_j\partial^\alpha\phi \circ T^{-sx})(y)|\,\mathrm{d}x &\le \int_0^1 \nu_{m,0}(\partial_j\partial^\alpha\phi \circ T^{-sx})\,\mathrm{d}s \le \int_0^1 (1+|sx|)^m \nu_{m,0}(\partial_j\partial^\alpha\phi)\,\mathrm{d}s \\ &\le (1+|x|)^m \nu_{m,0}(\partial_j\partial^\alpha\phi) \le (1+|x|)^m \nu_{m,|\alpha|+1}(\phi)\end{aligned}$$

by (B.12), we may apply (B.15) to $\partial^\alpha\phi$ instead of ϕ to conclude that for all $m, k \in \mathbf{Z}_{\ge 0}$ we have

$$\nu_{m,k}(\phi \circ T^{-x} - \phi) \le n|x|(1+|x|)^m \nu_{m,k+1}(\phi) \to 0 \quad \text{as} \quad x \to 0.$$

This proves the desired result.

2. For each $\phi \in \mathcal{S}(\mathbf{R}^n)$, $t \in \mathbf{R}$ and $j \in \{1, \dots, n\}$, we define $\Delta_t^j\phi \in \mathcal{S}(\mathbf{R}^n)$ by

$$\Delta_t^j \phi(x) := \frac{\phi(x+te_j) - \phi(x)}{t},$$

where e_j is the canonical j-th basis vector in $\mathbf{R}^n$. We claim that $\Delta_t^j\phi$ converges to $\partial_j\phi$ in $\mathcal{S}(\mathbf{R}^n)$ as $t \to 0$. Writing

$$\phi(x+he_j) - \phi(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\phi(x+ste_j)\,\mathrm{d}s = t\int_0^1 \partial_j\phi(x+ste_j)\,\mathrm{d}s,$$

we conclude that

$$\Delta_t^j \phi(x) - \partial_j \phi(x) = \int_0^1 ((\partial_j \phi \circ T^{-ste_j})(x) - \partial_j \phi(x)) \, \mathrm{d}s.$$

Hence, by applying this to $\partial^\alpha \phi$ instead of ϕ, for all $m, k \in \mathbf{Z}_{\geq 0}$ we have

$$\nu_{m,k}(\Delta_t^j \phi - \partial_j \phi) \leq \nu_{m,k+1}(\phi \circ T^{-ste_j} - \phi) \to 0 \quad \text{as} \quad s \to 0$$

by step 1. This proves the result.

3. Now let $u \in \mathcal{S}'(\mathbf{R}^n)$, $\phi \in \mathcal{S}(\mathbf{R}^n)$, $x \in \mathbf{R}^n$, and $j \in \{1, \ldots, n\}$. Then

$$\frac{(u * \phi)(x + te_j) - (u * \phi)(x)}{t} = u\left(\frac{\phi \circ T^{x+te_j} - \phi \circ T^x}{t}\right) = -u(\Delta_{-t}^j(\phi \circ T^x))$$

for all $t \in \mathbf{R}$, so that

$$\partial_j(u * \phi)(x) = \lim_{t \to 0} \frac{(u * \phi)(x + te_j) - (u * \phi)(x)}{t} = -u(\partial_j(\phi \circ T^x)) = (\partial_j u * \phi)(x)$$

by step 2 and Lemma B.39. Since

$$(\partial_j u * \phi)(x) = -u(\partial_j(\phi \circ T^x)) = u(\partial_j \phi * T^x) = (u * \partial_j \phi)(x),$$

we may proceed by induction to conclude that (B.14) holds.

Next we will show that $u * \phi \in \mathcal{O}_M(\mathbf{R}^n)$. Since $u \in \mathcal{S}'(\mathbf{R}^n)$, there are $m, k \in \mathbf{Z}_{\geq 0}$ and a $c \in \mathbf{R}_+$ so that $|u(\psi)| \leq c\nu_{m,k}(\psi)$ for all $\psi \in \mathcal{S}(\mathbf{R}^n)$. Hence, by (B.14) and (B.13), for all multi-indices α we have

$$\begin{aligned} |\partial^\alpha(u * \phi)(x)| = |(u * \partial^\alpha \phi)(x)| = |u(\partial^\alpha \phi \circ T^x)| &\leq c\nu_{m,k}(\partial^\alpha \phi \circ T^x) \\ &\leq c(1 + |x|)^m \nu_{m,k}(\partial^\alpha \phi) \leq c(1 + |x|)^m \nu_{m,k+|\alpha|}(\phi) \end{aligned}$$

so that $u * \phi \in \mathcal{O}_M(\mathbf{R}^n)$, as desired. Moreover, we note that this inequality implies that

$$n_{k',\psi}(u * \phi) \leq c \sup_{x \in \mathbf{R}^n} (1 + |x|)^m |\psi(x)| \nu_{m,k+k'}(\phi) = c\nu_{m,0}(\psi)\nu_{m,k+k'}(\phi)$$

so that the convolution operator $\mathcal{S}(\mathbf{R}^n) \to \mathcal{O}_M(\mathbf{R}^n)$, $\phi \mapsto u * \phi$ is continuous.

For the final assertion we refer to [Gr, Proposition 2.3.22(11)]. □

B.42 Definition. Let $P(D)$ be a linear partial differential operator with constant coefficients, i.e., an operator as in (B.11). We call $E \in \mathcal{D}'(\mathbf{R}^n)$ a *fundamental solution* of $P(D)$ if $P(D)E = \delta$. ♢

B.43 Proposition. *Let P be a polynomial and let $P(D)$ be the associated linear partial differential operator. Suppose $P(D)$ has a fundamental solution $E \in \mathcal{S}'(\mathbf{R}^n)$. Then the convolution operator*

$$\mathscr{L} : \mathcal{S}(\mathbf{R}^n) \to \mathcal{O}_M(\mathbf{R}^n), \quad \mathscr{L}\phi := E * \phi$$

satisfies

$$P(D)(\mathscr{L}\phi) = \phi \quad \textit{for all } \phi \in \mathcal{S}(\mathbf{R}^n).$$

In particular, for a given $\phi \in \mathcal{S}(\mathbf{R}^n)$ *the function* $u := \mathscr{L}\phi \in \mathcal{O}_M(\mathbf{R}^n)$ *is a solution to the partial differential equation* $P(D)u = \phi$. *Moreover, the Fourier transform of the fundamental solution satisfies the equation*

$$P(\xi)\mathscr{F}E = 1.$$

Proof. Let $\phi \in \mathcal{S}(\mathbf{R}^n)$. By linearity we have

$$P(D)(\mathscr{L}\phi) = P(D)(E * \phi) = P(D)E * \phi = \delta * \phi = \phi.$$

This proves the first assertion.

For the second assertion we note that

$$1 = \mathscr{F}\delta = \mathscr{F}P(D)E = P(\xi)\mathscr{F}E$$

by Example B.34. The assertion follows. □

C Appendix: Complex Analysis

Rather than differentiating with the coordinates obtained from $\mathbf{R}^2$, i.e., using the directional derivatives ∂_x and ∂_y, we will use the so-called *Wirtinger derivatives*

$$\partial_z := \frac{1}{2}(\partial_x - i\partial_y), \quad \partial_{\overline{z}} := \frac{1}{2}(\partial_x + i\partial_y). \tag{C.1}$$

To motivate why it makes sense to define these operators in this way, we write $z = x + iy$ and $\overline{z} = x - iy$. Then for a given C^1-function $f : \mathbf{R}^2 \to \mathbf{C}$ we can write

$$F(z, \overline{z}) := f\left(\frac{z + \overline{z}}{2}, \frac{z - \overline{z}}{2i}\right) = f(x, y).$$

Formally taking derivatives and using the chain rule yields

$$\begin{aligned} \partial_x f(x, y) &= \partial_z F(z, \overline{z}) + \partial_{\overline{z}} F(z, \overline{z}) \\ \partial_y f(x, y) &= i\partial_z F(z, \overline{z}) - i\partial_{\overline{z}} F(z, \overline{z}), \end{aligned}$$

which implies $\partial_x = \partial_z + \partial_{\overline{z}}$ and $i\partial_y = \partial_{\overline{z}} - \partial_z$. Solving for ∂_z and $\partial_{\overline{z}}$ yields (C.1).

We note in particular that $\partial_z z = \partial_{\overline{z}} \overline{z} = 1$ and $\partial_z \overline{z} = \partial_{\overline{z}} z = 0$.

C.1 Proposition. *Let $U \subseteq \mathbf{C}$ be open and let f, g be differentiable in U. Then*

(i) $\overline{\partial_z f} = \partial_{\overline{z}} \overline{f}$, *and* $\overline{\partial_{\overline{z}} f} = \partial_z \overline{f}$;

(ii) $\partial_z(fg) = g\partial_z f + f\partial_z g$, *and* $\partial_{\overline{z}}(fg) = g\partial_{\overline{z}} f + f\partial_{\overline{z}} g$;

(iii) If $g(U) \subseteq U$, then

$$\begin{aligned} \partial_z(f \circ g) &= (\partial_z f \circ g)\partial_z g + (\partial_{\overline{z}} f \circ g)\partial_z \overline{g} \\ \partial_{\overline{z}}(f \circ g) &= (\partial_{\overline{z}} f \circ g)\partial_{\overline{z}} g + (\partial_z f \circ g)\partial_{\overline{z}} \overline{g}. \end{aligned}$$

Let $a, b \in \mathbf{R}$ with $a < b$. We say that a function $\gamma : [a, b] \to \mathbf{C}$ is a C^1 path if it is continuous, continuously differentiable in $]a, b[$, and its derivative extends continuously to $[a, b]$. We say that such a function is a *piecewise* C^1 path if there is a partition $a = a_0 < a_1 < \cdots < a_m = b$ of $[a, b]$ so that the restriction $\gamma_j := \gamma|_{[a_{j-1}, a_j]}$ is of class C^1 for all $j \in \{1, \ldots, m\}$. We call $(\gamma_j)_{j=1}^m$ the C^1 pieces of γ. We say that γ is a closed path, if $\gamma(a) = \gamma(b)$.

C.2 Definition. Suppose $\gamma : [a, b] \to \mathbf{C}$ is a C^1 path and suppose $f : \gamma([a, b]) \to \mathbf{C}$ is a continuous function. Then we define the integral and the conjugate integral of f along γ by

$$\int_\gamma f(z)\,\mathrm{d}z := \int_a^b f(\gamma(t))\gamma'(t)\,\mathrm{d}t, \quad \int_\gamma f(z)\,\mathrm{d}\overline{z} := \int_a^b f(\gamma(t))\overline{\gamma'(t)}\,\mathrm{d}t,$$

respectively.

More generally, if γ is a piecewise C^1 path, then we may define the integral and the conjugate integral of f along γ by

$$\int_\gamma f(z)\,\mathrm{d}z := \sum_{j=1}^m \int_{\gamma_j} f(z)\,\mathrm{d}z, \quad \int_\gamma f(z)\,\mathrm{d}\overline{z} := \sum_{j=1}^m \int_{\gamma_j} f(z)\,\mathrm{d}\overline{z},$$

where $(\gamma_j)_{j=1}^m$ are the C^1 pieces of γ. ♢

We note that we have the relation

$$\int_\gamma f(z)\,\mathrm{d}\overline{z} = \overline{\int_\gamma \overline{f(z)}\,\mathrm{d}z}.$$

C.3 Lemma. *Suppose $\gamma : [a,b] \to \mathbf{C}$ is a C^1 path and suppose $f : \gamma([a,b]) \to \mathbf{C}$ is a continuous function. Let $a', b' \in \mathbf{R}$ with $a' < b'$. If $\phi : [a', b'] \to [a,b]$ is a C^1 mapping satisfying $\phi(a') = a$ and $\phi(b') = b$, then*

$$\int_\gamma f(z)\,\mathrm{d}z = \int_{\gamma\circ\phi} f(z)\,\mathrm{d}z$$

and similarly for the conjugate integral. If however $\phi(a') = b$ and $\phi(b') = a$, then

$$\int_\gamma f(z)\,\mathrm{d}z = -\int_{\gamma\circ\phi} f(z)\,\mathrm{d}z$$

and similarly for the conjugate integral.

Proof. Since $(\gamma\circ\phi)'(t) = \gamma'(\phi(t))\phi'(t)$ for $t \in]a', b'[$, we find by the Change of Variables Theorem that

$$\int_{\gamma\circ\phi} f(z)\,\mathrm{d}z = \int_{a'}^{b'} f(\gamma(\phi(t)))\gamma'(\phi(t))\phi'(t)\,\mathrm{d}t = \int_{\phi(a')}^{\phi(b')} f(\gamma(t))\gamma'(t)\,\mathrm{d}t.$$

The integral on the right is equal to $\int_\gamma f(z)\,\mathrm{d}z$ if $\phi(a') = a$ and $\phi(b') = b$ and to $-\int_\gamma f(z)\,\mathrm{d}z$ if $\phi(a') = b$ and $\phi(b') = a$, as desired. The proof for the conjugate integral is analogous. □

For a piecewise C^1 path $\gamma : [a,b] \to \mathbf{C}$ we define $\gamma^{-1} : [a,b] \to \mathbf{C}$ by $\gamma^{-1}(t) := \gamma(a+b-t)$. Then the above lemma implies that

$$\int_{\gamma^{-1}} f(z)\,\mathrm{d}z = -\int_\gamma f(z)\,\mathrm{d}z$$

for any continuous function f defined on $\gamma([a,b])$, and similarly for the conjugate integral.

C.4 Definition. A set $\Gamma \subseteq \mathbf{C}$ is called a *Jordan curve* if there is a continuous closed path $\gamma : [a,b] \to \mathbf{C}$ that is injective on $[a,b[$, satisfies $\gamma(a) = \gamma(b)$, and satisfies $\gamma([a,b]) = \Gamma$. In this case we say that Γ is parameterized by γ. If γ is a piecewise C^1 path with non-vanishing derivatives where it is C^1, then we call Γ a piecewise C^1 Jordan curve or a *closed contour.* ♢

Basic examples of piecewise C^1 Jordan curves are circles and rectangles in $\mathbf{C}$.

C.5 Theorem (Jordan Curve Theorem)**.** *The complement of any Jordan curve in $\mathbf{C}$ is the disjoint union of exactly two non-empty connected open sets.*

A proof may be found in [Ha, p. 169]. Since a Jordan Curve Γ is compact, precisely one of its complementing components must be bounded. We say that Γ *encases* the bounded component.

C.6 Definition. Let $\Omega \subseteq \mathbf{C}$ be a bounded open set so that its boundary $\partial\Omega$ consists of finitely many disjoint piecewise C^1 Jordan curves $\Gamma_1, \ldots \Gamma_m$. We say that $\partial\Omega$ is *positively oriented* if for each $j \in \{1, \ldots, m\}$ we have chosen a parameterization γ_j of Γ_j so that γ_j traverses Γ_j while leaving Ω to its left. For a continuous function $f : \partial\Omega \to \mathbf{C}$ we then define the *contour integrals*

$$\oint_{\partial\Omega} f(z)\,\mathrm{d}z := \sum_{j=1}^{m} \int_{\gamma_j} f(z)\,\mathrm{d}z, \quad \oint_{\partial\Omega} f(z)\,\mathrm{d}\overline{z} := \sum_{j=1}^{m} \int_{\gamma_j} f(z)\,\mathrm{d}\overline{z}.$$

$\diamondsuit$

The notion of a parameterization of a Jordan curve leaving a region to its left is well-defined by the Jordan Curve Theorem. We remark that the definition of the integral along $\partial\Omega$ as above is independent of the chosen parameterizations of the Jordan curves by Lemma C.3. Green's Integral Theorem takes on a particularly nice form when using our current notation.

C.7 Theorem (Green's Integral Theorem)**.** *Let $\Omega \subseteq \mathbf{C}$ be a bounded open set so that its boundary $\partial\Omega$ consists of finitely many disjoint piecewise C^1 Jordan curves. Let $f : \overline{\Omega} \to \mathbf{C}$ be a continuous function that is C^1 in Ω. Then*

$$\int_{\Omega} \partial_z f(z)\,\mathrm{d}z = -\frac{1}{2i} \oint_{\partial\Omega} f(z)\,\mathrm{d}\overline{z}, \quad \int_{\Omega} \partial_{\overline{z}} f(z)\,\mathrm{d}z = \frac{1}{2i} \oint_{\partial\Omega} f(z)\,\mathrm{d}z,$$

where the integrals over Ω are integrals with respect to the Lebesgue measure on Ω.

An immediate consequence of Green's Integral Theorem is Cauchy's Theorem.

C.8 Corollary (Cauchy's Theorem)**.** *Let $U \subseteq \mathbf{C}$ be open. Suppose $f \in C^1(U)$ satisfies $\partial_{\overline{z}} f(z) = 0$ for all $z \in \Omega$. Let $\Gamma \subseteq U$ be a piecewise C^1 Jordan curve parameterized by γ, that encases a subset of U. Then*

$$\int_{\gamma} f(z)\,\mathrm{d}z = 0.$$

Proof. If we denote the set encased by Γ by Ω, then $\overline{\Omega} \subseteq U$ and the result is immediate from Green's Integral Theorem. $\square$

C.9 Example. Let $r, R \in \mathbf{R}$ so that $0 < r < R$. Then we define the annulus $\Omega := \{z \in \mathbf{C} \mid r < |z| < R\}$. Then $\partial\Omega$ consists of the circles with radii r and R, centered at the origin. For $\rho \in \mathbf{R}_+$ we define

$$\gamma_\rho : [0, 2\pi] \to \mathbf{C}, \quad \gamma_\rho(t) := \rho e^{it}.$$

Then $\partial\Omega$ is given a positive orientation by the C^1 paths γ_r^{-1} and γ_R.

Let $f : \overline{\Omega} \to \mathbf{C}$ be a continuous function that is C^1 in Ω and satisfies $\partial_{\overline{z}} f(z) = 0$ for all $z \in \Omega$. Then Green's Integral Theorem implies that

$$0 = 2i \int_{\Omega} \partial_{\overline{z}} f(z)\,\mathrm{d}z = \int_{\gamma_R} f(z)\,\mathrm{d}z - \int_{\gamma_r} f(z)\,\mathrm{d}z.$$

Hence,

$$\int_{\gamma_R} f(z)\,\mathrm{d}z = \int_{\gamma_r} f(z)\,\mathrm{d}z,$$

meaning that, where it is defined, the integral of f along counterclockwise oriented circles around 0 is independent of the radius of the circle.

Moreover, we find that

$$\begin{aligned}\int_\Omega \partial_z f(z)\,\mathrm{d}z &= \frac{1}{2i}\int_{\gamma_r} f(z)\,\mathrm{d}\overline{z} - \frac{1}{2i}\int_{\gamma_R} f(z)\,\mathrm{d}\overline{z}\\ &= \frac{1}{2}\int_0^{2\pi} (Rf(Re^{it}) - rf(re^{it}))e^{-it}\,\mathrm{d}t.\end{aligned} \tag{C.2}$$

If $f(z) = -1/z$ for $z \in \partial\Omega$, then $\partial_z f(z) = 1/z^2$ for $z \in \Omega$, and we may conclude from (C.2) that

$$\int_\Omega \frac{1}{z^2}\,\mathrm{d}z = \frac{1}{2}\int_0^{2\pi} (e^{-it} - e^{-it})e^{-it}\,\mathrm{d}t = 0. \tag{C.3}$$

◇

C.10 Theorem (Cauchy-Pompeiu Integral Formula)**.** *Let $\Omega \subseteq \mathbf{C}$ be a bounded open set so that its boundary $\partial\Omega$ consists of finitely many disjoint piecewise C^1 Jordan curves. Let $f : \overline{\Omega} \to \mathbf{C}$ be a continuous function that is C^1 in Ω. Then*

$$f(w) = \frac{1}{2\pi i}\oint_{\partial\Omega} \frac{f(z)}{z-w}\,\mathrm{d}z - \frac{1}{\pi}\int_\Omega \frac{\partial_{\overline{z}} f(z)}{z-w}\,\mathrm{d}z$$

for all $w \in \Omega$.

Proof. Let $w \in \Omega$. Pick $\varepsilon \in \mathbf{R}_+$ so that the closed disc $\overline{D}_\varepsilon$ of radius ε centered at w lies in Ω. Then consider the open set $\Omega_\varepsilon := \Omega\backslash\overline{D}_\varepsilon$. By the product rule for $\partial_{\overline{z}}$ and the fact that $\partial_{\overline{z}} 1/(z-w) = 0$ in Ω_ε we find that $1/(w-z)\partial_{\overline{z}} f = \partial_{\overline{z}}(f/(z-w))$ in Ω_ε. We define $\gamma : [0, 2\pi] \to \mathbf{C}$ by $\gamma(t) := w + \varepsilon e^{it}$. Then Green's Integral Theorem implies that

$$\begin{aligned}\int_{\Omega_\varepsilon} \frac{\partial_{\overline{z}} f(z)}{z-w}\,\mathrm{d}z &= \frac{1}{2i}\oint_{\partial\Omega} \frac{f(z)}{z-w}\,\mathrm{d}z - \frac{1}{2i}\int_\gamma \frac{f(z)}{z-w}\,\mathrm{d}z\\ &= \frac{1}{2i}\oint_{\partial\Omega} \frac{f(z)}{z-w}\,\mathrm{d}z - \frac{1}{2i}\int_0^{2\pi} \frac{f(w+\varepsilon e^{it})}{\varepsilon e^{it}} i\varepsilon e^{it}\,\mathrm{d}t.\end{aligned} \tag{C.4}$$

Since

$$\lim_{\varepsilon\downarrow 0}\int_0^{2\pi} f(w+\varepsilon e^{it})\,\mathrm{d}t = 2\pi f(w),$$

and since $\frac{1}{z-w}$ is integrable over D_ε, we conclude from letting $\varepsilon \downarrow 0$ in (C.4) that

$$\int_\Omega \frac{\partial_{\overline{z}} f(z)}{z-w}\,\mathrm{d}z = \frac{1}{2i}\oint_{\partial\Omega} \frac{f(z)}{z-w}\,\mathrm{d}z - \pi f(w).$$

The assertion follows. □

An immediate consequence is the following:

C.11 Corollary (Cauchy's Integral Formula)**.** *Let $\Omega \subseteq \mathbf{C}$ be a bounded open set so that its boundary $\partial\Omega$ consists of finitely many disjoint piecewise C^1 Jordan curves. Let $f : \overline{\Omega} \to \mathbf{C}$ be a continuous function that is C^1 in Ω and satisfies $\partial_{\overline{z}} f(z) = 0$ for all $z \in \Omega$. Then*

$$f(w) = \frac{1}{2\pi i} \oint_{\partial\Omega} \frac{f(z)}{z - w} \,\mathrm{d}z.$$

for all $w \in \Omega$.

We can also show the following:

C.12 Corollary. *The locally integrable function $E := 1/(\pi z)$ is a fundamental solution of $\partial_{\overline{z}}$, i.e., $\partial_{\overline{z}} E = \delta$ in $\mathcal{D}'(\mathbf{C})$.*

Proof. Fix $\phi \in C_c^\infty(\mathbf{C})$. Pick $R \in \mathbf{R}_+$ large enough so that $\phi(z) = 0$ whenever $|z| \geq R$. Let Ω be the open disc of radius R around the origin. Then, since $\oint_{\partial\Omega} \phi(z)/z \,\mathrm{d}z = 0$, it follows from the Cauchy-Pompeiu Integral Formula that

$$\langle \partial_{\overline{z}} E, \phi \rangle = -\frac{1}{\pi} \int_\Omega \frac{\partial_{\overline{z}} \phi(z)}{z} \,\mathrm{d}z = \phi(0) = \langle \delta, \phi \rangle.$$

We conclude that $\partial_{\overline{z}} E = \delta$ in $\mathcal{D}'(\mathbf{C})$, as asserted. □

C.13 Definition. Let $U \subseteq \mathbf{C}$ be open and let $f : U \to \mathbf{C}$ be a function. We say that f is *complex differentiable* at $z_0 \in U$ if

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

exists. If this is the case, then we denote the limit by $f'(z_0)$. If f is complex differentiable at all points in U, then we say that f is *holomorphic* in U with (complex) derivative $f' : U \to \mathbf{C}$. ♢

C.14 Lemma. *Let $U \subseteq \mathbf{C}$ be open. A function $f : U \to \mathbf{C}$ is complex differentiable at $z_0 \in U$ if and only if it is (totally) differentiable in the sense of differentiability for maps from the real two-dimensional vector space $\mathbf{C}$ to itself at z_0 and satisfies $\partial_{\overline{z}} f(z_0) = 0$. Moreover, in this case we have $\partial_z f(z_0) = f'(z_0)$.*

Proof. Suppose f is complex differentiable at z_0. Left multiplication by $f'(z_0) = a + bi$, $a, b \in \mathbf{R}$, is represented by the matrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

If we denote the corresponding linear map by L, then

$$\lim_{z \to z_0} \frac{|f(z) - f(z_0) - L(z - z_0)|}{|z - z_0|} = \lim_{z \to z_0} \left| \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \right| = 0,$$

meaning that $f = u + iv$ is (totally) differentiable at z_0 with (total) derivative given by

$$\begin{pmatrix} \partial_x u(z_0) & \partial_y u(z_0) \\ \partial_x v(z_0) & \partial_y v(z_0) \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

But this implies that we have $\partial_x u(z_0) = \partial_y v(z_0)$ and $\partial_y u(z_0) = -\partial_x v(z_0)$, or equivalently,

$$\partial_{\overline{z}} f(z_0) = \frac{1}{2}(\partial_x u(z_0) - \partial_y v(z_0) + i(\partial_y u(z_0) + \partial_x v(z_0))) = 0.$$

We note that this also implies that

$$\partial_z f(z_0) = \frac{1}{2}(\partial_x u(z_0) + \partial_y v(z_0) + i(\partial_x v(z_0) - \partial_y u(z_0))) = \partial_x u(z_0) + i\partial_y u(z_0) = a + bi = f'(z_0).$$

For the converse, suppose $f = u + iv$ is continuously differentiable at z_0 and satisfies $\partial_{\overline{z}} f(z_0) = 0$. Then we find that $\partial_x u(z_0) = \partial_y v(z_0)$ and $\partial_y u(z_0) = -\partial_x v(z_0)$. As above, this implies that $\partial_z f(z_0) = \partial_x u(z_0) + i\partial_y u(z_0)$. Moreover, it implies that the derivative of f is given by the Jacobian matrix

$$L := \begin{pmatrix} \partial_x u(z_0) & \partial_y u(z_0) \\ \partial_x v(z_0) & \partial_y v(z_0) \end{pmatrix} = \begin{pmatrix} \partial_x u(z_0) & \partial_y u(z_0) \\ -\partial_y u(z_0) & \partial_x u(z_0) \end{pmatrix}.$$

But then

$$\lim_{z \to z_0} \left| \frac{f(z) - f(z_0)}{z - z_0} - \partial_z f(z_0) \right| = \lim_{z \to z_0} \frac{|f(z) - f(z_0) - L(z - z_0)|}{|z - z_0|} = 0.$$

The assertion follows. □

Using the chain rule, one can now verify the following:

C.15 Lemma. *Let $U \subseteq \mathbf{C}$ be open and let $f \in C^1(U)$. Suppose $\gamma : [a, b] \to \mathbf{C}$ is a C^1 path whose image lies in U. Then*

$$(f \circ \gamma)'(t) = \partial_z f(\gamma(t))\gamma'(t) + \partial_{\overline{z}} f(\gamma(t))\overline{\gamma'(t)}$$

for all $t \in]a, b[$. In particular, if f is holomorphic, then $(f \circ \gamma)'(t) = f'(\gamma(t))\gamma'(t)$ for all $t \in]a, b[$.

Let $(a_j)_{j \in \mathbf{Z}_{\geq 0}}$ be a sequence of complex numbers. Then we define $r \in [0, \infty]$ by $1/r = \limsup_{j \to \infty} |a_j|^{1/j}$, which we call the convergence radius of this sequence. This terminology is justified by the fact that the power series

$$\sum_{j \in \mathbf{Z}_{\geq 0}} a_j z^j$$

converges absolutely for $z \in \mathbf{C}$ with $|z| < r$, uniformly for $z \in \mathbf{C}$ with $|z| \leq r'$ where $r' < r$, and diverges whenever $|z| > r$. We sometimes call r the convergence radius of the corresponding power series.

C.16 Definition. Let $U \subseteq \mathbf{C}$ be open and let $f : U \to \mathbf{C}$ be a function. We say that f is *analytic* at $z_0 \in U$ if there is a sequence $(a_j)_{j \in \mathbf{Z}_{\geq 0}}$ with positive convergence radius r so that

$$f(z) = \sum_{j \in \mathbf{Z}_{\geq 0}} a_j (z - z_0)^j \tag{C.5}$$

for $|z - z_0| < r$. If f is analytic at all points in U, then we say that f is analytic in U. ◇

A typical example is the exponential function $\exp : \mathbf{C} \to \mathbf{C}$ defined by

$$\exp(z) := e^z := \sum_{j \in \mathbf{Z}_{\geq 0}} \frac{z^j}{j!}.$$

The radius of convergence of this series is ∞. In the same way as is done for the real case, one can show that $e^{z+w} = e^z e^w$ for all $z, w \in \mathbf{C}$. Moreover, we have

$$e^{it} = \sum_{j \in \mathbf{Z}_{\geq 0}} \frac{(it)^j}{j!} = \sum_{j \in \mathbf{Z}_{\geq 0}} (-1)^j \frac{t^{2j}}{(2j)!} + i \sum_{j \in \mathbf{Z}_{\geq 0}} (-1)^j \frac{t^{2j+1}}{(2j+1)!} = \cos t + i \sin t$$

for all $t \in \mathbf{R}$.

Suppose f is as in the definition and satisfies (C.5). Formally differentiating term by term would indicate that

$$f'(z) = \sum_{j \in \mathbf{Z}_{\geq 0}} (j+1) a_{j+1} (z - z_0)^j$$

for $|z - z_0| < r$. This is justified by the fact that $\limsup_{j \to \infty} |(j+1)a_{j+1}|^{1/j} = \limsup_{j \to \infty} |a_j|^{1/j} = 1/r$. In particular, if f is analytic in U, then f is holomorphic in U, and its complex derivative is again analytic. In view of Lemma C.14, this means that we may conclude that analytic functions are smooth. It is actually true that any holomorphic function is analytic.

C.17 Theorem. *Let $U \subseteq \mathbf{C}$ be open and let $f : U \to \mathbf{C}$ be a function. The following are equivalent:*

(i) f is holomorphic in U;

(ii) $f \in C^1(U)$ and $\partial_{\overline{z}} f(z) = 0$ for all $z \in U$;

(iii) f is analytic in U.

C.18 Remark. The operator $\partial_{\overline{z}}$ is an elliptic partial differential operator with constant coefficients. The elliptic regularity theory tells us that if a distribution $u \in \mathcal{D}'(U)$ satisfies $\partial_{\overline{z}} u = 0$, then u is given by a smooth function. The result of the theorem then implies that this function must be holomorphic. In particular, condition (ii) in the theorem may be replaced by the equivalent condition

(ii') f is continuous and

$$\int_U f(z) \partial_{\overline{z}} \phi(z) \, \mathrm{d}z = 0$$

for all $\phi \in C_c^\infty(U)$, i.e., $\partial_{\overline{z}} f = 0$ in $\mathcal{D}'(U)$.

$\diamondsuit$

For the proof of Theorem C.17, we note that we have already established the implications (iii)$\Rightarrow$(ii) in the discussion preceding the theorem and (ii)$\Rightarrow$(i) in Lemma C.14. Tt remains to show the implication (i)$\Rightarrow$(iii). This follows from Proposition C.21 below. For this proposition, we require a result known as Goursat's Theorem. While this result is a version of Cauchy's Theorem, it is stronger in the sense that continuity of the derivatives of the function is not required. By a rectangle, we mean a closed rectangle whose sides are parallel to the coordinate axes in the plane. Per convention, we give its boundary a counterclockwise (positive) orientation.

C.19 Theorem (Goursat's Theorem)**.** *Let $U \subseteq \mathbf{C}$ be open and let $f : U \to \mathbf{C}$ be holomorphic. Suppose $R \subseteq U$ is a closed rectangle. Then*

$$\oint_{\partial R} f(z)\,\mathrm{d}z = 0.$$

Proof. We claim that we can find rectangles

$$R \supseteq R_1 \supseteq R_2 \supseteq R_3 \supseteq \cdots$$

so that, if L_j denotes the length of the boundary of R_j and L the length of the boundary of R and d_j denotes the diameter of R_j and d The diameter of R, we have the properties

$$\frac{1}{4^j}\left|\oint_{\partial R} f(z)\,\mathrm{d}z\right| \leq \left|\oint_{\partial R_j} f(z)\,\mathrm{d}z\right|, \quad L_j = \frac{1}{2^j}L, \quad d_j = \frac{1}{2^j}d \tag{C.6}$$

for all $j \in \mathbf{N}$. Indeed, we bisect the sides of R to subdivide it into four rectangles R^1, R^2, R^3, R^4. If one traverses the boundary of each of these rectangles counterclockwise, one notes that the sides of the rectangle in the interior of R are traversed once forwards and once backwards so that

$$\int_{\partial R} f(z)\,\mathrm{d}z = \sum_{k=1}^{4} \int_{\partial R^k} f(z)\,\mathrm{d}z$$

and thus

$$\left|\int_{\partial R} f(z)\,\mathrm{d}z\right| \leq \sum_{k=1}^{4}\left|\int_{\partial R^k} f(z)\,\mathrm{d}z\right| \leq 4 \max_{k\in\{1,2,3,4\}}\left|\int_{\partial R^k} f(z)\,\mathrm{d}z\right|$$

Pick $k \in \{1,2,3,4\}$ where this maximum is attained and set $R_1 := R^k$. Then $L_1 = L/2$, $d_1 = d/2$ and (C.6) holds for $j = 1$.

Now suppose we have shown that we can pick subrectangles $R_1, \dots, R_J$ of R so that (C.6) holds for $j \in \{1, \dots, J\}$ for some $J \in \mathbf{N}$. Using the same bisection process from before on R_J, we obtain a rectangle R_{J+1} so that $L_{J+1} = L_J/2 = L/2^{J+1}$, $d_{J+1} = d_J/2 = d/2^{J+1}$ and

$$\frac{1}{4^J}\left|\oint_{\partial R} f(z)\,\mathrm{d}z\right| \leq \left|\oint_{\partial R_J} f(z)\,\mathrm{d}z\right| \leq 4\left|\int_{\partial R^{J+1}} f(z)\,\mathrm{d}z\right|,$$

which proves the induction step. This proves the claim.

By Cantor's Intersection Theorem for complete metric spaces we find that the intersection $\bigcap_{j\in\mathbf{N}} R_j$ consists of a single point z_0. Since f is complex differentiable at z_0, we find that the function

$$r : U\backslash\{z_0\} \to \mathbf{C}, \quad r(z) := \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0)$$

continuously extends to z_0, where it takes the value 0. Now note that for each $j \in \mathbf{N}$ we have

$$\begin{aligned}\oint_{\partial R_j} f(z)\,\mathrm{d}z &= f(z_0)\oint_{\partial R_j} \mathrm{d}z + f'(z_0)\oint_{\partial R_j}(z - z_0)\,\mathrm{d}z + \oint_{\partial R_j}(z - z_0)r(z)\,\mathrm{d}z \\ &= \oint_{\partial R_j}(z - z_0)r(z)\,\mathrm{d}z,\end{aligned}$$

where the first two integrals vanish by Cauchy's Theorem. Then (C.6) gives us the estimate

$$\frac{1}{4^j}\left|\oint_{\partial R} f(z)\,\mathrm{d}z\right| \le \left|\oint_{\partial R_j} f(z)\,\mathrm{d}z\right| \le \oint_{\partial R_j} |z-z_0||r(z)|\,\mathrm{d}z \le d_j L_j \sup_{z\in R_j}|r(z)| = \frac{dL}{4^j}\sup_{z\in R_j}|r(z)|.$$

Hence,

$$\left|\oint_{\partial R} f(z)\,\mathrm{d}z\right| \le dL\sup_{z\in R_j}|r(z)| \to 0 \quad \text{as} \quad j\to\infty.$$

The assertion follows. □

C.20 Corollary. *Let $\Omega \subseteq \mathbf{C}$ be an open disc centered at $z_0 \in C$ and let $f : \Omega \to \mathbf{C}$ be holomorphic. Then there is a holomorphic function $F : \Omega \to \mathbf{C}$ so that $F' = f$.*

Proof. For $w \in \Omega$ we set

$$F(w) := \int_{z_0}^{w} f(z)\,\mathrm{d}z,$$

which should be interpreted as the integral of f along a path following the sides of the rectangle R with opposing vertices z_0 and w. To see that this is well-defined, we first note that there are two possible choices of such paths which we shall call γ_1 and γ_2. If one traverses γ_1 followed by traversing γ_2 backwards, we note that we have traversed the boundary of R where we assume without loss of generality that this has been done counterclockwise. Hence, by Goursat's Theorem, we have

$$\int_{\gamma_1} f(z)\,\mathrm{d}z - \int_{\gamma_2} f(z)\,\mathrm{d}z = \oint_{\partial R} f(z)\,\mathrm{d}z = 0.$$

This proves that F is well-defined.

Pick $h \in \mathbf{C}$. Then, by making appropriate choices of traversed sides of rectangles, we note that

$$F(w+h) - F(w) = \int_{w}^{w+h} f(z)\,\mathrm{d}z,$$

where the integral is taken along a path following the sides of a rectangle R_h with opposing vertices w and $w+h$. We note that the length of this path can be estimated by $|\operatorname{Re} h| + |\operatorname{Im} h|$. Moreover, we note that

$$F(w+h) - F(w) = \int_{w}^{w+h} (f(z) - f(w) + f(w))\,\mathrm{d}z = hf(w) + \int_{w}^{w+h} (f(z) - f(w))\,\mathrm{d}z. \tag{C.7}$$

Since f is continuous at w, we find that

$$\left|\frac{1}{h}\int_{w}^{w+h} (f(z)-f(w))\,\mathrm{d}z\right| \le \frac{|\operatorname{Re} h| + |\operatorname{Im} h|}{|h|}\sup_{z\in\partial R_h}|f(z)-f(w)| \le \sqrt{2}\sup_{z\in\partial R_h}|f(z)-f(w)| \to 0$$

as $h \to 0$. Thus, we conclude from (C.7) that

$$\lim_{h\to 0}\frac{F(w+h)-F(w)}{h} = f(w).$$

This proves the desired result. □

C.21 Proposition. *Let $U \subseteq \mathbf{C}$ be open and let $z_0 \in U$ and $R \in \mathbf{R}_+$ so that the closure in $\mathbf{C}$ of the open disc Ω of radius R centered at z_0 lies in U. Suppose a function $f : U \to \mathbf{C}$ is holomorphic in U. Then*

$$f(w) = \frac{1}{2\pi i} \oint_{\partial\Omega} \frac{f(z)}{z - w} \,\mathrm{d}z$$

for all $w \in \Omega$. Moreover, f is analytic at z_0, and its series representation has a convergence radius greater than or equal to R.

Proof. Fix $w \in \Omega$. First, we define

$$g : U\backslash\{w\} \to \mathbf{C}, \quad g(z) := \frac{f(z) - f(w)}{z - w}.$$

Then g is holomorphic in $U\backslash\{w\}$. For $\rho \in \mathbf{R}_+$ small enough so that the closed disc centered at w of radius ρ lies in Ω we define

$$\gamma_\rho : [0, 2\pi] \to \mathbf{C}, \quad \gamma_\rho(t) := w + \rho e^{it}.$$

Moreover, we set $\gamma_R : [0, 2\pi] \to \mathbf{C}$, $\gamma_R(t) := z_0 + e^{it}$ and we fix a $\rho \in \mathbf{R}_+$ as before. Consider a partition $0 = a_0 \le a_1 \le \cdots \le a_J = 2\pi$ of the interval $[0, 2\pi]$ so that there exist discs $\{\Omega_j\}_{j=0}^{J-1}$ in $U\backslash\{w\}$ so that

$$\gamma_R([a_j, a_{j+1}]) \subseteq \Omega_j, \quad \gamma_\rho([a_j, a_{j+1}]) \subseteq \Omega_j$$

for all $j \in \{0, \ldots, J-1\}$. We write $v_j := \gamma_R(a_j)$, $w_j := \gamma_\rho(a_j)$ for all $j \in \{0, \ldots, J\}$. By Corollary C.20 we can find holomorphic primitives $G_j : \Omega_j \to \mathbf{C}$ of the restriction of g to Ω_j for all $j \in \{0, \ldots, J-1\}$. Note that for each $z \in \Omega_j \cap \Omega_{j+1}$ we have

$$\partial_z(G_j - G_{j+1})(z) = g(z) - g(z) = 0, \quad \partial_{\overline{z}}(G_j - G_{j+1})(z) = 0.$$

This implies that $G_j - G_{j+1}$ is constant in $\Omega_j \cap \Omega_{j+1}$. In particular, since $v_{j+1}, w_{j+1} \in \Omega_j \cap \Omega_{j+1}$, we have

$$G_{j+1}(v_{j+1}) - G_{j+1}(w_{j+1}) = G_j(v_{j+1}) - G_j(w_{j+1}) \quad \text{for all } j \in \{0, \ldots, J-2\}. \tag{C.8}$$

Then, by the Fundamental Theorem of Calculus

$$\int_{\gamma_\rho} g(z) \,\mathrm{d}z = \sum_{j=0}^{J-1} \int_{a_j}^{a_{j+1}} G_j'(\gamma_\rho(t))\gamma_\rho'(t) \,\mathrm{d}t = \sum_{j=0}^{J-1} \left(G_j(w_{j+1}) - G_j(w_j)\right)$$

Similarly we have

$$\int_{\gamma_R} g(z) \,\mathrm{d}z = \sum_{j=0}^{J-1} \left(G_j(v_{j+1}) - G_j(v_j)\right).$$

Hence, by (C.8), we have

$$\begin{aligned}
\int_{\gamma_R} g(z) \,\mathrm{d}z - \int_{\gamma_\rho} g(z) \,\mathrm{d}z &= \sum_{j=0}^{J-1} \left(G_j(v_{j+1}) - G_j(w_{j+1}) - (G_j(v_j) - G_j(w_j))\right) \\
&= G_{J-1}(v_J) - G_{J-1}(w_J) - (G_0(v_0) - G_0(w_0)) \\
&= G_{J-1}(z_0) - G_0(z_0) - (G_{J-1}(w_0) - G_0(z_0)) = 0,
\end{aligned}$$

since $v_0 = z_0 + R = v_J$, $w_0 = w + \rho = w_J$ and $G_{J-1} - G_0$ is constant on the intersection of their domains. Thus, we have shown that the integral of g along γ_ρ does not depend on ρ. Now, since f is complex differentiable at z_0, g extends continuously to z_0 and is thus bounded in a neighborhood of z_0 by a constant $c \in \mathbf{R}_+$. This implies that for small enough $\rho \in \mathbf{R}_+$ we have

$$\left| \int_{\gamma_\rho} g(z)\,\mathrm{d}z \right| \leq c \int_0^{2\pi} |\gamma_\rho'(t)|\,\mathrm{d}t = 2\pi\rho c,$$

for which the right-hand side tends to 0 as $\rho \downarrow 0$. Thus, we have shown that

$$0 = \int_{\gamma_R} g(z)\,\mathrm{d}z = \int_{\gamma_R} \frac{f(z)}{z-w}\,\mathrm{d}z - f(w) \int_{\gamma_R} \frac{1}{z-w}\,\mathrm{d}z$$

so that

$$\oint_{\partial\Omega} \frac{f(z)}{z-w}\,\mathrm{d}z = f(w) \int_{\gamma_R} \frac{1}{z-w}\,\mathrm{d}z = f(w) \int_0^{2\pi} \frac{iRe^{it}}{Re^{it}}\,\mathrm{d}t = 2\pi i f(w).$$

This proves the first assertion.

For the last assertion, let $0 < \rho < R$. Then, since

$$\left| \frac{w-z_0}{z-z_0} \right| = \frac{|w-z_0|}{R} < 1$$

for all $w \in \Omega$ and $z \in \partial\Omega$, we find that

$$\frac{1}{z-w} = \frac{1}{z-z_0} \frac{1}{1 - \frac{w-z_0}{z-z_0}} = \frac{1}{z-z_0} \sum_{j \in \mathbf{Z}_{\geq 0}} \left(\frac{w-z_0}{z-z_0} \right)^j$$

for all $w \in \Omega$ and $z \in \partial\Omega$, where the sum is absolutely and uniformly convergent for $|w - z_0| \leq \rho$. Then dominated convergence implies that

$$f(w) = \sum_{j \in \mathbf{Z}_{\geq 0}} \left(\frac{1}{2\pi i} \oint_{\partial\Omega} \frac{f(z)}{(z-z_0)^{j+1}}\,\mathrm{d}z \right) (w-z_0)^j,$$

where the sum is absolutely and uniformly convergent for $|w - z_0| \leq \rho$. The assertion follows. □

We note that we may now use the terms analytic and holomorphic interchangeably. In particular, we obtain the following:

C.22 Corollary. *Let $U \subseteq \mathbf{C}$ be open. If $f : U \to \mathbf{C}$ is holomorphic in U, then f' is also holomorphic in U. Thus, f is infinitely many times complex differentiable in U.*

We present a converse to Cauchy's Theorem. We say that an open connected set $U \subseteq \mathbf{C}$ is *simply connected* if all Jordan curves in U encase a subset of U. As an example we note that open discs are simply connected, whereas open annuli are not. As an extreme case of this, removing a point from $\mathbf{C}$ results in a connected open set that is not simply connected.

C.23 Theorem (Morera's Theorem)**.** *Let $U \subseteq \mathbf{C}$ be open. Suppose a continuous function $f : U \to \mathbf{C}$ satisfies*

$$\oint_\Gamma f(z)\,\mathrm{d}z = 0 \tag{C.9}$$

for all piecewise C^1 Jordan curves $\Gamma \subseteq U$ encasing a subset of U. Then, on each simply connected open subset $V \subseteq U$, there is a holomorphic function $F : V \to \mathbf{C}$ satisfying $F'(z) = f(z)$ for all $z \in V$. In particular, f is holomorphic in U and, if U itself is simply connected, then f admits a holomorphic primitive in U.

Proof. Let $z_0 \in U$ and let V be a simply connected open subset of U containing z_0, e.g., a disc of small enough radius centered at z_0. Then for each $w \in V \backslash \{z_0\}$ there exists an injective piecewise C^1 path $\gamma : [0,1] \to \mathbf{C}$ in V so that $\gamma(0) = z_0$, $\gamma(1) = w$. Let $\tilde{\gamma}$ be another injective piecewise C^1 path in V from z_0 to w. Then the concatenated path ψ which first traverses γ and then traverses $\tilde{\gamma}^{-1}$ describes a piecewise C^1 Jordan curve. Then (C.9) implies that

$$\int_\gamma f(z)\,\mathrm{d}z - \int_{\tilde{\gamma}} f(z)\,\mathrm{d}z = \int_\psi f(z)\,\mathrm{d}z = 0.$$

Thus, we may define $F : V \to \mathbf{C}$ by

$$F(w) := \int_{z_0}^{w} f(z)\,\mathrm{d}z := \int_\gamma f(z)\,\mathrm{d}z.$$

Let $\varepsilon \in \mathbf{R}_+$. And let $w \in V$. By continuity of f at w we can find a $\delta \in \mathbf{R}_+$ so that $w' \in V$ and $|f(w) - f(w')| < \varepsilon$ whenever $|w - w'| < \delta$. Suppose $w' \in V\backslash\{w\}$ satisfies $|w - w'| < \delta$. Define $\gamma : [0,1] \to \mathbf{C}$ by $\gamma(t) = (1-t)w' + tw$. Then

$$F(w) - F(w') = \int_\gamma f(z)\,\mathrm{d}z = (w - w')f(w) + \int_\gamma (f(z) - f(w))\,\mathrm{d}z. \tag{C.10}$$

Since $\gamma([0,1])$ lies in the ball centered at w of radius δ, we have

$$\frac{1}{|w-w'|}\left|\int_\gamma (f(z) - f(w))\,\mathrm{d}z\right| \le \varepsilon \frac{1}{|w-w'|}\int_0^1 |\gamma'(t)|\,\mathrm{d}t = \varepsilon.$$

We conclude from (C.10) that

$$\lim_{w'\to w} \frac{F(w) - F(w')}{w - w'} = f(w) + \lim_{w' \to w} \frac{1}{w-w'}\int_\gamma (f(z) - f(w))\,\mathrm{d}z = f(w).$$

The assertion follows. □

An important example is the definition of the complex logarithm.

C.24 Example. Let U be the complement in $\mathbf{C}$ of the non-positive real numbers. Then U is a simply connected open set in $\mathbf{C}$. Consider the holomorphic function $f : U \to \mathbf{C}$ defined by $f(z) := 1/z$. Then, by Cauchy's Theorem, f satisfies (C.9) for all piecewise C^1 Jordan curves

$\Gamma \subseteq U$ encasing a subset of U. The proof of Morera's Theorem shows that we may define a holomorphic primitive F of f on U by

$$\log(w) := \int_1^w f(z)\,\mathrm{d}z,$$

where the integral should be interpreted as the integral from 1 to w along a piecewise C^1 path from 1 to w. Note that F coincides with the natural logarithm on the positive real numbers. Moreover, we note that

$$\partial_z(ze^{-\log z}) = e^{-\log z} - e^{-\log z} = 0.$$

Since also $\partial_{\bar{z}}(ze^{-\log z}) = 0$ we conclude that $\partial_x(ze^{-\log z}) = \partial_y(ze^{-\log z}) = 0$ and thus, since U is connected, $ze^{-\log z} = c$ for some $c \in \mathbf{C}$. Taking $z = 1$ shows that $c = 1$ and

$$e^{\log z} = z \quad \text{for all } z \in U.$$

Note that we can write any $z \in U$ as $re^{i\alpha}$ with $r \in \mathbf{R}_+$ and $\alpha \in]-\pi, \pi[$. Define $\gamma : [0, \alpha] \to \mathbf{C}$ by $\gamma(t) := re^{it}$. Then

$$\log(re^{it}) = \int_1^r f(z)\,\mathrm{d}z + \int_\gamma f(z)\,\mathrm{d}z = \log r + \int_0^\alpha i\,\mathrm{d}t = \log r + i\alpha.$$

Next, we wish to define the logarithm of a function. Suppose $V \subseteq \mathbf{C}$ is a simply connected open set and suppose $f : V \to \mathbf{C}$ is holomorphic and satisfies $f(z) \neq 0$ for all $z \in V$. Let z_0 be any point in V and pick a point $w_0 \in \mathbf{C}$ so that $e^{w_0} = f(z_0)$. Then we define a holomorphic function $L_f : V \to \mathbf{C}$ by

$$L_f(w) := w_0 + \int_{z_0}^w \frac{f'(z)}{f(z)}\,\mathrm{d}z.$$

A proof analogous to the one above shows that $e^{L_f(z)} = f(z)$ for all $z \in V$. ♢

C.25 Definition. Let $z, w \in \mathbf{C}\backslash\{0\}$. We define the *oriented angle from z to w* to be the unique number $\alpha \in [0, 2\pi[$ so that

$$\frac{w}{|w|} = e^{i\alpha}\frac{z}{|z|}.$$

Let $U \subseteq \mathbf{C}$ be open and let $z_0 \in U$. If $\gamma_j : [a_j, b_j] \to \mathbf{C}$, $j \in \{1, 2\}$ are two C^1 paths in U so that $\gamma_j(t_j) = z_0$ for some $t_j \in]a_j, b_j[$, $j \in \{1, 2\}$, then we say that γ_1 and γ_2 are *paths through z_0*. If $\gamma_j'(t_j) \neq 0$ for $j \in \{1, 2\}$, then there is a well-defined oriented angle $\alpha \in [0, 2\pi[$ from $\gamma_1'(t_1)$ to $\gamma_2'(t_2)$. We define the *oriented angle from γ_1 to γ_2 at z_0* to be α.

A C^1 function $f : U \to \mathbf{C}$ is said to *preserve oriented angles at z_0* if the following conditions hold: if γ_1 and γ_2 are paths through z_0 whose oriented angle $\alpha \in [0, 2\pi[$ at z_0 from γ_1 to γ_2 is well defined, then the oriented angle from $f \circ \gamma_1$ to $f \circ \gamma_2$ at $f(z_0)$ is also well-defined and is equal to α. The function f is called *conformal* if it preserves oriented angles at all points in U. ♢

In the notation of this definition, note that if f is holomorphic, then Lemma C.15 implies that $(f \circ \gamma_j)'(t_j) = f'(z_0)\gamma_j'(t_j)$ for $j \in \{1, 2\}$. Thus, for the oriented angle from $f \circ \gamma_1$ to $f \circ \gamma_2$ at $f(z_0)$ to be defined, it is necessary that $f'(z_0) \neq 0$. In fact, this is also a sufficient condition for preservation of angles.

C.26 Proposition. *Let $U \subseteq \mathbf{C}$ be open and let $f : U \to \mathbf{C}$ be a holomorphic function. Then f preserves oriented angles at $z_0 \in U$ if and only if $f'(z_0) \neq 0$.*

Proof. Assume $f'(z_0) \neq 0$. Let γ_1 and γ_2 be paths through z_0 in U whose oriented angle $\alpha \in [0, 2\pi[$ at z_0 from γ_1 to γ_2 is well defined. If $\gamma_1(t_1) = \gamma_2(t_2) = z_0$ we write $z_1 := \gamma_1'(t_1) \neq 0$, $z_2 := \gamma_2'(t_2) \neq 0$. Then $(f \circ \gamma_j)'(t_j) = f'(z_0)z_j \neq 0$ for $j \in \{1, 2\}$ so that the oriented angle from $f \circ \gamma_1$ to $f \circ \gamma_2$ at $f(z_0)$ is also well-defined. Let $\alpha \in [0, 2\pi[$ be the oriented angle from γ_1 to γ_2 at z_0. Then,

$$\frac{(f \circ \gamma_2)'(t_2)}{|(f \circ \gamma_2)'(t_2)|} \Big/ \frac{(f \circ \gamma_1)'(t_1)}{|(f \circ \gamma_1)'(t_1)|} = \left(\frac{f'(z_0)}{|f'(z_0)|}\frac{z_2}{|z_2|}\right) \Big/ \left(\frac{f'(z_0)}{|f'(z_0)|}\frac{z_1}{|z_1|}\right) = \frac{z_2}{|z_2|} \Big/ \frac{z_1}{|z_1|} = e^{i\alpha}.$$

The assertion follows.

□

Next we present an Inverse Function Theorem for holomorphic functions as a consequence of the Inverse Function Theorem from real analysis.

C.27 Definition. Let $U \subseteq \mathbf{C}$ be open and $f : U \to \mathbf{C}$ a holomorphic function. We say that f is a *biholomorphism* if $V := f(U)$ is open and there is a holomorphic map $g : V \to C$ satisfying $g(V) = U$, $g(f(z)) = z$ for all $z \in U$ and $f(g(z)) = z$ for all $z \in V$. We say that f is *biholomorphic* at a point z_0 if there is an open neighborhood $U' \subseteq U$ of z_0 so that the restriction of f to U' is a biholomorphism. ♢

C.28 Theorem (Inverse Function Theorem)**.** *Let $U \subseteq \mathbf{C}$ be open and $f : U \to \mathbf{C}$ a holomorphic function. If $f'(z_0) \neq 0$ for some $z_0 \in U$, then f is biholomorphic at z_0.*

Proof. Recall the proof of Lemma C.14. Write $f'(z_0) = a + bi$. Then the Jacobian matrix of f at z_0 is given by

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

This has determinant $a^2 + b^2 = |f'(z_0)|^2 \neq 0$. Thus, the real analysis Inverse Function Theorem implies that there is an open neighborhood $U' \subseteq \mathbf{C}$ of z_0 in U and an open set $V \subseteq \mathbf{C}$ so that $f(U') = V$, and a function $g \in C^1(V)$ so that $g(V) = U'$, that inverts f. The chain rule implies that the Jacobian matrix of g at $f(z_0)$ is given by the inverse matrix

$$\frac{1}{|f'(z_0)|^2}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

But from this it follows that

$$\partial_{\bar{z}} g(f(z_0)) = \frac{1}{2|f'(z_0)|}((a - a) + i(-b + b) = 0.$$

Similarly, one shows that $\partial_{\bar{z}} g(w) = 0$ for all $w \in V$. Hence, g is holomorphic in V, as asserted. □

The converse to the preceding theorem, that if a function is biholomorphic at a point, then its derivative at that point is non-zero, also holds. We actually have a stronger result:

C.29 Proposition. *Let $U \subseteq \mathbf{C}$ be open and $f : U \to \mathbf{C}$ a holomorphic function. If f is injective, then $f'(z) \neq 0$ for all $z \in U$ and f is a biholomorphism.*

For the proof we require a lemma.

C.30 Lemma. *Let $U \subseteq \mathbf{C}$ be open and let $f : U \to \mathbf{C}$ be a holomorphic function. Suppose the series expansion at $z_0 \in U$ of f is given by*

$$f(z) = a_0 + \sum_{j \in \mathbf{Z}_{\geq k}} a_j (z - z_0)^j,$$

for some $k \in \mathbf{Z}_{\geq 1}$ with $a_k \neq 0$. Then there is an open neighborhood $U' \subseteq U$ of z_0 and a holomorphic function $g : U' \to \mathbf{C}$ so that g is biholomorphic at z_0 and

$$f(z) = a_0 + g(z)^k$$

for all $z \in U'$.

Proof. We may assume $a_0 = 0$ and $z_0 = 0$ by considering $z \mapsto f(z + z_0) - f(z_0)$ instead of f. Then, for $|z|$ small enough, define h so that

$$f(z) = a_k z^k \sum_{j \in \mathbf{Z}_{\geq 0}} \frac{a_{j+k}}{a_k} z^j = a_k z^k h(z).$$

Then h is holomorphic and, since $h(0) = 1$, it is non-zero in an open neighborhood V of 0. Let $L_h : V \to \mathbf{C}$ denote the holomorphic logarithm of h as defined in Example C.24. Pick $s \in \mathbf{C}$ so that $s^k = a_k$. Then we may define a holomorphic function $g : V \to \mathbf{C}$ by

$$g(z) := s z e^{\frac{L_h(z)}{k}}.$$

Then $g(z)^k = f(z)$ for all $z \in V$ and $g'(0) = s e^{L_h(0)/k} \neq 0$ by the assumption $s^k = a_k \neq 0$. By the Inverse Function Theorem, g is biholomorphic at 0. The assertion follows. □

Proof of Proposition C.29. The function $U \to f(U)$ given by f is bijective, hence has an inverse function h. In particular, note that injectivity of f implies that f is nowhere constant so that we may apply Lemma C.30. Let $z_0 \in U$ and let g and k be as in the lemma so that $f(z) = f(z_0) + g(z)^k$ for z in an open neighborhood U' of z_0. Injectivity of f on U' implies that $z \mapsto \phi(z) := g(z + z_0)^k$ must be injective in a neighborhood of 0. But since g is biholomorphic at z_0, this means that $w \mapsto w^k$ is injective in a neighborhood of 0. This is only possible when $k = 1$. But this means that $f'(z_0) = a_1 \neq 0$, which implies that f is biholomorphic at z_0. As z_0 was arbitrary, this means that f must be locally invertible by a holomorphic map at every point in U. But since each of these inverses must be given by restrictions of h, we may conclude that h is holomorphic. The assertion follows. □

We can now characterize biholomorphisms.

C.31 Theorem. *Let $U \subseteq \mathbf{C}$ be open and let $f \in C^1(U)$. The following are equivalent:*

(i) f is a biholomorphism;

(ii) f is injective and holomorphic;

(iii) f is injective and conformal.

For the proof we require a lemma.

C.32 Lemma. *Let $U \subseteq \mathbf{C}$ be open, $f \in C^1(U)$, and $z_0 \in U$. Then if f preserves angles at z_0, then $\partial_{\bar{z}} f(z_0) = 0$. In particular, if f is conformal, then f is holomorphic.*

Proof. For $\theta \in \mathbf{R}$ we define a C^1 path γ_θ by $\gamma_\theta(t) := z_0 + te^{-i\theta/2}$ on an interval around 0, small enough so that the image of γ_θ lies in U. Then $\gamma_\theta(0) = z_0$ and $\gamma_\theta'(0) = e^{-i\theta/2} \neq 0$ so that

$$\frac{\overline{\gamma_\theta'(0)}}{\gamma_\theta'(0)} = e^{i\theta}. \tag{C.11}$$

Then, since f preserves angles at z_0, we have for $\alpha, \beta \in \mathbf{R}$ that

$$\frac{(f \circ \gamma_\alpha)'(0)}{|(f \circ \gamma_\alpha)'(0)|} \Big/ \frac{(f \circ \gamma_\beta)'(0)}{|(f \circ \gamma_\beta)'(0)|} = \frac{\gamma_\alpha'(0)}{|\gamma_\alpha'(0)|} \Big/ \frac{\gamma_\beta'(0)}{|\gamma_\beta'(0)|},$$

or equivalently,

$$\frac{|\gamma_\alpha'(0)|}{|(f \circ \gamma_\alpha)'(0)|} \frac{(f \circ \gamma_\alpha)'(0)}{\gamma_\alpha'(0)} = \frac{|\gamma_\beta'(0)|}{|(f \circ \gamma_\beta)'(0)|} \frac{(f \circ \gamma_\beta)'(0)}{\gamma_\beta'(0)}.$$

This means that $(f \circ \gamma_\alpha)'(0)/\gamma_\alpha'(0)$ and $(f \circ \gamma_\beta)'(0)/\gamma_\beta'(0)$ have the same argument. But then, since α, β were arbitrary, we find, by Lemma C.15 and (C.11), that the argument of

$$\frac{(f \circ \gamma_\theta)'(0)}{\gamma_\theta'(0)} = \partial_z f(z_0) + \partial_{\bar{z}} f(z_0) e^{i\theta} \tag{C.12}$$

is independent of $\theta \in \mathbf{R}$. Since (C.12) describes a circle of radius $|\partial_{\bar{z}} f(z_0)|$ as θ runs through $[0, 2\pi]$, we conclude that we must have $\partial_{\bar{z}} f(z_0) = 0$. The result follows. The last assertion follows from Theorem C.17. □

Proof of Theorem C.31. The equivalence of (i) and (ii) follows from Proposition C.29. The implication (ii)⇒(iii) follows from Proposition C.26 while the implication (iii)⇒(ii) follows from Lemma C.32. □

We can also prove the Open Mapping Theorem for holomorphic functions.

C.33 Theorem (Open Mapping Theorem)**.** *Let $U \subseteq \mathbf{C}$ be open. Suppose $f : U \to \mathbf{C}$ is a holomorphic function that is non-constant on any non-empty open subset of U. Then f maps open sets to open sets.*

Proof. Let $V \subseteq U$ be open and let $z_0 \in V$ so that $f(z_0)$ is an arbitrary point in $f(V)$. We have to show that there is an open neighborhood of $f(z_0)$ that is contained in $f(V)$.

By Lemma C.30 we can find an open neighborhood $U' \subseteq V$ of z_0 and a holomorphic function $g : U' \to \mathbf{C}$ so that g is biholomorphic at z_0 and

$$f(z) = f(z_0) + g(z)^k$$

for all $z \in U'$. As g is a biholomorphism on an open neighborhood $W \subseteq U'$ of z_0, the set $g(W)$ is open. Since $g(z_0) = 0$, there is an open disc D around the origin contained in $g(W)$. Since the disc D gets mapped to another disc D' through the map $w \mapsto w^k$, we conclude that

$$f(g^{-1}(D)) = f(z_0) + D'$$

is an open neighborhood of $f(z_0)$ contained in $f(V)$. We conclude that $f(V)$ is open, as desired. □

We define

$$\Delta := -4\partial_z\partial_{\overline{z}} = -\partial_x^2 - \partial_y^2$$

and recall that for an open set $U \subseteq \mathbf{C}$ a real-valued function $u \in C^2(U)$ is called *harmonic* if $\Delta u = 0$ in U.

C.34 Proposition. *Let $U \subseteq \mathbf{C}$ be a simply connected open set and suppose a real-valued $u \in C^2(U)$ is harmonic. Then there is a holomorphic function $f : U \to \mathbf{C}$ so that* $\operatorname{Re} f = u$*. Conversely, the real part of any holomorphic function is harmonic.*

Proof. Set $g := 2\partial_z u \in C^1(U)$. Then $\partial_{\overline{z}} g(z) = -\Delta u(z)/2 = 0$ for all $z \in U$. Thus, g is holomorphic in U. By Morera's Theorem, g has a holomorphic primitive $f : U \to \mathbf{C}$. Set $\tilde{u} := \operatorname{Re} f$. Then, recalling the proof of Lemma C.14,

$$\partial_x u(z) - i\partial_y u(z) = g(z) = f'(z) = \partial_z f(z) = \partial_x \tilde{u}(z) - i\partial_y \tilde{u}(z)$$

for all $z \in U$. This implies that $u - \tilde{u} \in C^1(U)$ has vanishing partial derivatives, hence must be equal to a constant $c \in \mathbf{R}$. Thus, u is the real part of the holomorphic function $f + c$. This proves the first assertion.

For the converse, let $V \subseteq \mathbf{C}$ be open and let $f : V \to \mathbf{C}$ be holomorphic. Then f is smooth, and certainly $u := \operatorname{Re} f \in C^2(V)$. Since $\partial_{\overline{z}} f(z) = 0$ for all $z \in V$, we have $\Delta u(z) = \operatorname{Re}(-4\partial_z\partial_{\overline{z}} f(z)) = 0$ for all $z \in V$. The result follows. □

For $z_0 \in \mathbf{C}$ and $r \in \mathbf{R}_+$ we will denote by $D(z_0; r)$ and $\overline{D}(z_0; r)$ the respectively open and closed disc of radius r in $\mathbf{C}$ centered at z_0. We denote the boundary of such a disc by $\partial D(z_0; r)$.

C.35 Corollary. *Let $V, U \subseteq \mathbf{C}$ be open and let $u : V \to \mathbf{R}$ be harmonic. If $f : U \to \mathbf{C}$ is a holomorphic function satisfying $f(U) \subseteq V$, then the composition $u \circ f : U \to \mathbf{R}$ is harmonic in U.*

Proof. If f is constant then the result is trivial, so assume f is non-constant. Let $z_0 \in U$. Pick $r \in \mathbf{R}_+$ so that $D(z_0; r) \subseteq U$. As $f(D(z_0; r))$ is open by the Open Mapping Theorem, there is some $r' \in \mathbf{R}_+$ so that $D(f(z_0); r') \subseteq f(D(z_0; r)) \subseteq V$. Since $D(f(z_0); r')$ is simply connected, there is a holomorphic function $g : D(z_0; r) \to \mathbf{C}$ so that $\operatorname{Re} g(z) = u(z)$ for all $z \in D(f(z_0); r')$. It follows that

$$u \circ f(z) = \operatorname{Re}((g \circ f)(z))$$

for all $z \in D(z_0; r)$. As the composition of holomorphic functions is holomorphic, we conclude that $u \circ f$ is holomorphic at z_0. Since z_0 was arbitrary, we conclude that $u \circ f$ is harmonic in U.

Alternatively, one can compute $\Delta(u \circ f)(z) = \Delta u(f(z))|f'(z)|^2 = 0$ for all $z \in U$. □

If $\overline{D}(z_0; r) \subseteq U$ for some $z_0 \in U$ and $r \in \mathbf{R}_+$, then u is the real part of some holomorphic function f defined on an open set in U containing $\overline{D}(z_0; r)$. If $\gamma : [0, 2\pi] \to \mathbf{C}$ is defined by $\gamma(t) := z_0 + re^{it}$, then we find, by Cauchy's integral formula, that

$$u(z_0) = \operatorname{Re} f(z_0) = \operatorname{Re} \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z_0 - z} \, \mathrm{d}z = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it}) \, \mathrm{d}t.$$

This motivates the following definition:

C.36 Definition. Let $U \subseteq \mathbf{C}$ be open and let $u : U \to \mathbf{R}$ be a continuous function. We call u *subharmonic* if

$$u(z_0) \leq \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it}) \, \mathrm{d}t$$

for all $z_0 \in U$, $r \in \mathbf{R}_+$ so that $\overline{D}(z_0; r) \subseteq U$. ♢

For real-valued functions $u \in C^2(U)$, subharmonicity is a condition on Δu.

C.37 Proposition. *Let $U \subseteq \mathbf{C}$ be open and let $u \in C^2(U)$ be real-valued. Then u is subharmonic in U if and only if $\Delta u \leq 0$ in U.*

Proof. Suppose $u \in C^2(U)$. Let $z_0 \in U$ and pick $r_0 \in \mathbf{R}_+$ so that $D(z_0; r_0) \subseteq U$. Define $\phi : [0; r_0[\to \mathbf{R}$ by

$$\phi(r) := \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it}) \, \mathrm{d}t.$$

For $r \in]0, r_0[$, define $\gamma_r : [0, 2\pi] \to \mathbf{C}$ by $\gamma_r(t) := z_0 + re^{it}$. Then, by Green's Integral Theorem, we have

$$-\frac{1}{2r} \int_{D(z_0;r)} \Delta u(z) \, \mathrm{d}z = \frac{2}{r} \int_{D(z_0;r)} \partial_{\overline{z}} \partial_z u(z) \, \mathrm{d}z = \frac{1}{ri} \int_\gamma \partial_z u(z) \, \mathrm{d}z = \int_0^{2\pi} \partial_z u(z_0 + re^{it}) e^{it} \, \mathrm{d}t$$

and

$$-\frac{1}{2r} \int_{D(z_0;r)} \Delta u(z) \, \mathrm{d}z = \frac{2}{r} \int_{D(z_0;r)} \partial_z \partial_{\overline{z}} u(z) \, \mathrm{d}z = -\frac{1}{ri} \int_\gamma \partial_{\overline{z}} u(z) \, \mathrm{d}\overline{z} = \int_0^{2\pi} \partial_z u(z_0 + re^{-it}) e^{-it} \, \mathrm{d}t,$$

so that

$$\phi'(r) = \frac{1}{2\pi} \int_0^{2\pi} \left(\partial_z u(z_0 + re^{it}) e^{it} + \partial_{\overline{z}} u(z_0 + re^{it}) e^{-it} \right) \, \mathrm{d}t = -\frac{1}{2\pi r} \int_{D(z_0;r)} \Delta u(z) \, \mathrm{d}z. \qquad \text{(C.13)}$$

Now suppose $\Delta u \leq 0$ in U. Then (C.13) implies that $\phi'(r) \geq 0$ for $r \in]0, r_0[$ so that ϕ is increasing. This implies that

$$u(z_0) = \phi(0) \leq \phi(r) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it}) \, \mathrm{d}t$$

for all $r \in]0, r_0[$. We conclude that u is subharmonic.

For the converse we argue by contraposition. Suppose that $\Delta u(z_0) > 0$ for some $z_0 \in U$. Then, by continuity of Δu, there is some $r_0 \in \mathbf{R}_+$ so that $D(z_0; r_0) \subseteq U$ and $\Delta u > 0$ on $D(z_0; r)$. Then, defining ϕ as before, by (C.13), we have $\phi'(r) < 0$ for $r \in]0, r_0[$. Hence, ϕ is strictly decreasing. This implies that

$$u(z_0) = \phi(0) > \phi(r) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it}) \, \mathrm{d}t$$

for $r \in]0, r_0[$. Thus, u is not subharmonic in U. The assertion follows. □

The notion of subharmonicity can be characterized in various ways.

C.38 Theorem. *Let $U \subseteq \mathbf{C}$ be open and let $u : U \to \mathbf{R}$ be a continuous function. The following are equivalent:*

(i) u is subharmonic in U;

(ii) for every $z_0 \in U$ there is an $r_0 \in \mathbf{R}_+$ so that $\overline{D}(z_0; r_0) \subseteq U$ and whenever $0 < r < r_0$, we have

$$u(z_0) \leq \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it}) \, \mathrm{d}t;$$

(iii) for every $z_0 \in U$ and $r \in \mathbf{R}_+$ so that $\overline{D}(z_0; r) \subseteq U$ we have that if a continuous function $v : \overline{D}(z_0; r) \to \mathbf{R}$, twice continuously differentiable and harmonic in $D(z_0; r)$, satisfies $u(z) \leq v(z)$ for all $z \in \partial D(z_0; r)$, then $u(z) \leq v(z)$ for all $z \in \overline{D}(z_0; r)$.

We will give a proof of this result momentarily. The implication (ii)⇒(iii) uses the so-called Maximum Principle, while the implication (iii)⇒(i) uses the fact that for all $z_0 \in \mathbf{C}$, $r \in \mathbf{R}_+$ and continuous $g : \partial D(z_0; r) \to \mathbf{R}$, there exists a unique continuous function $u : \overline{D}(z_0; r) \to \mathbf{R}$, twice continuously differentiable in $D(z_0; r)$, solving the Dirichlet problem

$$\begin{cases} \Delta u = 0 & \text{in } D(z_0; r); \\ u|_{\partial D(z_0; r)} = g. \end{cases}$$

A proof of this result can be found in [Ev, Chapter 2, Theorem 15].

C.39 Proposition (The Maximum Principle)**.** *Let $U \subseteq \mathbf{C}$ be open and connected and let $u : \overline{U} \to \mathbf{R}$ be continuous. Suppose that for every $z_0 \in U$ there is an $r_0 \in \mathbf{R}_+$ so that $\overline{D}(z_0; r_0) \subseteq U$ and whenever $0 < r < r_0$, we have*

$$u(z_0) \leq \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + re^{it}) \, \mathrm{d}t. \tag{C.14}$$

If for some $z_0 \in U$ we have $u(z_0) = \max_{z \in \overline{U}} u(z)$, then u is constant in $\overline{U}$.

Proof. Set $M := \max_{z\in\overline{U}} u(z)$. We let W be the subset of U consisting of all $z \in U$ so that $u(z) = M$. Per assumption, W is non-empty. By continuity of u, W is closed in U. We will show that W is also open in U. Then it follows from connectedness of U that $U = W$, as desired. Let $z_0 \in W$ and pick $R \in \mathbf{R}_+$ so that $\overline{D}(z_0; R) \subseteq U$. Per assumption we can find $0 < r_0 < R$ so that u satisfies (C.14) for any $r \in \mathbf{R}_+$ satisfying $0 < r < r_0$. Fix $0 < \varepsilon < r_0$. By employing polar coordinates we find that

$$\frac{1}{\pi\varepsilon^2}\int_{D(z_0;\varepsilon)} u(z)\,\mathrm{d}z = \frac{1}{\pi\varepsilon^2}\int_0^{\varepsilon}\int_0^{2\pi} u(z_0+\rho e^{it})\rho\,\mathrm{d}t\,\mathrm{d}\rho \geq \frac{2\pi u(z_0)}{\pi\varepsilon^2}\int_0^{\varepsilon}\rho\,\mathrm{d}\rho = M$$

so that

$$M \leq \frac{1}{\pi\varepsilon^2}\int_{D(z_0;\varepsilon)} u(z)\,\mathrm{d}z \leq \frac{M}{\pi\varepsilon^2}\int_{D(z_0;\varepsilon)}\mathrm{d}z = M.$$

But this is only possible if $u(z) = M$ for all $z \in D(z_0; \varepsilon)$. We conclude that $D(z_0; \varepsilon) \subseteq W$, proving that W is open. It now follows that u is constant in U. By continuity this means that u is constant in $\overline{U}$, as desired. □

Proof of Theorem C.38. The implication (i)⇒(ii) is clear.

For (ii)⇒(iii), pick $z_0 \in U$ and $r \in \mathbf{R}_+$ so that $\overline{D}(z_0; r) \subseteq U$. Suppose $v : \overline{D}(z_0; r) \to \mathbf{R}$ is a continuous function, twice continuously differentiable and harmonic in $D(z_0; r)$, and satisfies $u(z) \leq v(z)$ for all $z \in \partial D(z_0; r)$. If we set $w := u - v$, then $w(z) \leq 0$ for all $z \in \partial D(z_0; r)$. Pick $z_1 \in \overline{D}(z_0; r)$ so that $w(z_1) = \max_{z\in\overline{D}(z_0;r)} w(z)$. We consider two cases.

First we assume that $z_1 \in \partial D(z_0; r)$. Then it follows that $u(z) - v(z) = w(z) \leq 0$ for all $z \in \overline{D}(z_0; r)$, as desired.

Next, we assume that $z_1 \in D(z_0; r)$. Since v is harmonic in $D(z_0; r)$, this means, in particular, that $-v$ is subharmonic in $D(z_0; r)$, and thus so is w as the sum of subharmonic functions. Hence, w satisfies the conditions for the Maximum Principle in $D(z_0; r)$ and is thus constant in $\overline{D}(z_0; r)$. Picking a $z' \in \partial D(z_0; r)$, we find that $u(z) - v(z) = w(z) = w(z') \leq 0$ for all $z \in \overline{D}(z_0; r)$. The assertion follows.

For (iii)⇒(i), let $z_0 \in U$, $r \in \mathbf{R}_+$ so that $\overline{D}(z_0; r) \subseteq U$. Let $v : \overline{D}(z_0; r) \to \mathbf{R}$ be the unique solution to the Dirichlet problem

$$\begin{cases} \Delta v = 0 & \text{in } D(z_0; r); \\ v|_{\partial D(z_0;r)} = u|_{\partial D(z_0;r)}. \end{cases}$$

Then $u(z) \leq v(z)$ for all $z \in \overline{D}(z_0; r)$. Thus, we have

$$u(z_0) \leq v(z_0) = \frac{1}{2\pi}\int_0^{2\pi} v(z_0 + re^{it})\,\mathrm{d}t = \frac{1}{2\pi}\int_0^{2\pi} u(z_0 + re^{it})\,\mathrm{d}t,$$

as desired. □

C.40 Lemma. *Let $V, U \subseteq \mathbf{C}$ be open and let $u : V \to \mathbf{R}$ be subharmonic. If $f : U \to \mathbf{C}$ is a holomorphic function satisfying $f(U) \subseteq V$, then the composition $u \circ f : U \to \mathbf{R}$ is subharmonic in U.*

Proof. Let $z_0 \in U$. First assume $f'(z_0) \neq 0$. Then it follows from the Inverse Function Theorem that there is an open neighborhood $U' \subseteq U$ of z_0 so that $V' := f(U')$ is open and there is a holomorphic map $g : V' \to C$ satisfying $g(V') = U'$, $g(f(z)) = z$ for all $z \in U'$ and $f(g(z)) = z$ for all $z \in V'$.

Let $r_0 \in \mathbf{R}_+$ so that $\overline{D}(z_0; r_0) \subseteq U'$ and let $0 < r < r_0$. Let $v : \overline{D}(z_0; r) \to \mathbf{R}$ be a continuous function that is twice continuously differentiable and harmonic in $D(z_0; r)$ and satisfies $u(f(z)) \leq v(z)$ for all $z \in \partial D(z_0; r)$. We have to show that $u(f(z)) \leq v(z)$ for all $z \in \overline{D}(z_0; r)$. Note that $f(\partial D(z_0; r))$ is a C^1 Jordan curve which encases the connected open set $f(D(z_0; r))$. Set $w := u - v \circ g : f(\overline{D}(z_0; r)) \to \mathbf{R}$. Our assumption implies that $w(z) \leq 0$ for all $z \in f(\partial D(z_0; r))$. Pick $z_1 \in f(\overline{D}(z_0; r))$ so that $w(z_1) = \max_{z \in \overline{D}(z_0; r)} w(z)$. We consider two cases.

First we assume that $z_1 \in \partial f(D(z_0; r)) = f(\partial D(z_0; r))$. Then it follows that $u(z) - v(g(z)) = w(z) \leq 0$ for all $z \in \overline{f(D(z_0; r))} = f(\overline{D}(z_0; r))$, as desired.

Next, we assume that $z_1 \in f(D(z_0; r))$. Since $v \circ g$ is harmonic in $f(D(z_0; r))$, this means, in particular, that $-v \circ g$ is subharmonic in $f(D(z_0; r))$, and thus so is w as the sum of subharmonic functions. Hence, w satisfies the conditions for the Maximum Principle in $f(D(z_0; r))$ and is thus constant in $f(\overline{D}(z_0; r))$. Picking a $z' \in f(\partial D(z_0; r))$, we find that $u(z) - v(g(z)) = w(z) = w(z') \leq 0$ for all $z \in f(\overline{D}(z_0; r))$.

As in the proof of Theorem C.38, it now follows that

$$u(f(z_0)) \leq \frac{1}{2\pi} \int_0^{2\pi} u(f(z_0 + re^{it})) \, \mathrm{d}t$$

for $0 < r < r_0$.

It remains to check the case when $f'(z_0) = 0$. If f is constant, then the result is clear. If not, then Lemma C.30 implies that there is an open neighborhood $U' \subseteq U$ of z_0 and a holomorphic function $g : U' \to \mathbf{C}$ so that g is biholomorphic at z_0 and

$$f(z) = f(z_0) + g(z)^k$$

for all $z \in U'$, for some $k \in \mathbf{Z}_{\geq 2}$. Let $z_0 \in U'' \subseteq U'$ so that g is a biholomorphism on U''. Now, since $g(z_0) = 0$, we can pick $0 < r_0' < 1$ small enough so that $D(0; r_0') \subseteq g(U'')$. Let $0 < r' < r_0'$. Then

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} u(f(z_0) + r'^k e^{ikt}) \, \mathrm{d}t &= \frac{1}{2k\pi} \int_0^{2k\pi} u(f(z_0) + r'^k e^{it}) \, \mathrm{d}t \\ &= \frac{k}{2k\pi} \int_0^{2\pi} u(f(z_0) + r'^k e^{it}) \, \mathrm{d}t \geq u(f(z_0)), \end{aligned}$$

since u is subharmonic at $f(z_0)$. But this means that $z \mapsto u(f(z_0) + z^k)$ is subharmonic at 0. Then, since $g'(z_0) \neq 0$ and g is holomorphic, we may use our previous result to conclude that $z \mapsto u(f(z_0) + g(z)^k) = u(f(z))$ is subharmonic at z_0, i.e., there is some $r_0 \in \mathbf{R}_+$ so that

$$u(f(z_0)) \leq \frac{1}{2\pi} \int_0^{2\pi} u(f(z_0 + re^{it})) \, \mathrm{d}t$$

for $0 < r < r_0$. The assertion follows.

□

Index of notation

References

[Al] L. V. Ahlfors. *Conformality with respect to Riemannian Metrics*. Ann. Acad. Sci. Fenn. A I 206 (1955).

[Al2] L. V. Ahlfors. *Lectures on Quasiconformal Mappings*. Van Nostrand, Princeton, NJ, 1966.

[AB] L. V. Ahlfors, L. Bers. *Riemann Mapping's Theorem for Variable Metrics*. Ann. of Math. (2) 72 (1960) 385-404.

[Am] H. Amann. *Vector Valued Distributions and Fourier Multipliers*. Zürich, 2003.

[As] K. Astala. *Area Distortion of Quasiconformal Mappings*. Acta Math. 173 (1994), no. 1, 37-60.

[AIM] K. Astala, T. Iwaniec, G. Martin. *Elliptic Partial Differential Equations and Quasiconformal Mappings in the Plane*. Princeton Mathematical Series. Princeton University Press, Princeton and Oxford, 2009.

[AIPS] K. Astala, T. Iwaniec, I. Prause, E. Saksman. *Burkholder Integrals, Morrey's Problem and Quasiconformal Mappings*. J. Amer. Math. Soc. 25 (2012), 507-531.

[Ba] J. M. Ball. *Does rank-one convexity imply quasiconvexity?* In Metastability and Incompletely Posed Problems, volume 3, pages 17-32. IMA volumes in Mathematics and its Applications, 1987.

[BM] J. M. Ball, F. Murat. *$W^{1,p}$-Quasiconvexity and Variational Problems for Multiple Integrals*. J. Functional Analysis, 58 (1984), 225-253.

[BJ] R. Bañuelos, P. Janakiraman. *L^p-bounds for the Beurling-Ahlfors Transform*. Amer. Math. Soc. 360 (2008), no. 7, 3603-3612.

[Be] A. Beurling. *The Collected Works of Arne Beurling*. Volume 1, Birkhäuser, Boston, 1989.

[Bo] B. V. Bojarski. *Generalized Solutions of a System of First Order Differential Equations of Elliptic Type with Discontinuous Coefficients*. Mat. Sb. 43 (85) (1957), 451-503.

[Bu] D. L. Burkholder. *Boundary Value Problems and Sharp Inequalities for Martingale Transforms*. Ann. Probab. Volume 12, Number 3 (1984), 647-702.

[Bu2] D. L. Burkholder. *Sharp Inequalities for Martingales and Stochastic Integrals*. Colloque Paul Lévy sur les Processus Stochastique (Palaiseau, 1987). Astérisque No. 157-158 (1988), 75-94.

[Da] B. Dacorogna. *Direct Methods in the Calculus of Variations*. Second edition. Applied Mathematical Sciences 78. Springer. 2008.

[DV] O. Dragičević, A. Volberg. *Bellman Functions, Littlewood-Paley Estimates and Asymptotics for the Ahlfors-Beurling Operator in $L^p(\mathbf{C})$*. Indiana University Mathematics Journal 54 (4) (2005), 971-995.

[DK] J. J. Duistermaat, J. A. C. Kolk. *Distributions: Theory and Applications.* Translated by J.P. van Braam Houckgeest. Cornerstones. Birkhäuser, New York, 2010.

[Ev] L. C. Evans. *Partial Differential Equations: Second Edition.* Graduate Studies in Mathematics, Volume 19. American Mathematical Society, Providence, Rhode Island, 2010.

[Ev2] L. C. Evans. *Weak Convergence Methods for Nonlinear Partial Differential Equations.* Conference Board of the Mathematical Sciences, Regional Conference Series in Mathematics 74. American Mathematical Society, Providence, Rhode Island, 1990.

[Go] W. Goldstein. *The Exponent of Integrability of Generalized Derivatives of Quasiconformal Homeomorphisms on the Plane.* Dokl. Akad. Nauk SSSR 250 (1980), 18-21.

[GR] F. W. Gehring, E. Reich. *Area Distortion under Quasiconformal Mappings.* Ann. Acad. Sci. Fenn. Ser. A. I. Math., 388 (1966), 3-15.

[Gr] L. Grafakos. *Classical Fourier Analysis.* Graduate Texts in Mathematics 249. Springer. 2008.

[Gru] G. Grubb. *Distributions and Operators.* Graduate Texts in Mathematics 252. Springer. 2009.

[Ha] A. Hatcher. *Algebraic Topology.* Cambridge University Press. 2001.

[Iw] T. Iwaniec. *Extremal Inequalities In Sobolev Spaces and Quasiconformal Mappings.* Z. Anal. Anwendungen 1 (1982), no. 6, 1-16.

[Iw2] T. Iwaniec. *Nonlinear Cauchy-Riemann Operators in* $\mathbf{R}^n$. Trans. Amer. Math. Soc., 354 (2002), 1961-1995.

[IM] T. Iwaniec. G. Martin. *Riesz Transform and Related Singular Integrals.* J. Reine Angew. Math 437 (1996), 25-57.

[La] S. Lang. *Complex Analysis.* Fourth edition. Graduate Texts in Mathematics. Springer. 1999.

[LZ] P. D. Lax, L. Zalcman. *Complex Proofs of Real Theorems.* University Lectures Series, Volume 58. American Mathematical Society, Providence, Rhode Island, 2012.

[Le] O. Lehto. *Remarks on the Integrability of the Derivatives of Quasiconformal Mappings.* Ann. Sci. Fenn. Series AI Math. 371 (1965), 3-8.

[Mo] C. B. Morrey. *Quasi-convexity and the Lower Semicontinuity of Multiple Integrals.* Pacific J. Math., 2 (1952), 25-53.

[Mo2] C. B. Morrey. *Multiple Integrals in the Calculus of Variations.* Springer. Berlin, 1966.

[Ni] B. Nieraeth. *The Spectrum of the Dirichlet Laplacian.* Bachelor's Thesis Mathematics. Utrecht University, 2014. http://dspace.library.uu.nl/handle/1874/297489

[Pi] S. Pichorides. *On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov.* Studia Mathematica 44.2 (1972), 165-179.

[Ri] M. Riesz. *Sur Les Fonctions Cunjuguées*. Mathematische Zeitschrift 27 (1928), 218-244.

[Šv] V. Šverák. *Rank-one convexity does not imply quasiconvexity*. Proc. Roy. Soc. Edinburbh Sect. A 120 (1992), 185-189.

[Ve] I. N. Vekua. *The Problem of Reduction to the Canonical Form of Differential Forms of Elliptic Type and the Generalized Cauchy-Riemann System*. Dokl. Akad. Nauk SSSR (N.S.) 100, (1995), 197-200.